GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:37 ; Search time 27.25 Seconds
(without alignments)
4014.622 Million cell updates/sec

Title: US-09-902-481B-3
Perfect score: 5879
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5839	99.3	1153	1	RWHU1B	cell surface glyco
2	4447	75.6	1153	2	S00551	leukocyte surface
3	3456	58.8	1163	1	RWHU1C	cell surface glyco
4	1532.5	26.1	1170	2	S03308	cell surface glyco
5	1516.5	25.8	1163	2	I56126	lymphocyte fuction
6	1128	19.2	1179	2	A53213	integrin alpha-E c
7	1085.5	18.5	1151	2	A45226	integrin alpha-1 c
8	1071	18.2	1170	2	I45914	integrin alpha 2 s
9	1057	18.0	1178	2	S44142	VLA-2 protein homo
10	1054	17.9	1181	2	A33998	integrin alpha-2 c
11	1049	17.8	1180	2	A35854	integrin alpha-1 c
12	666	11.3	1039	2	A41131	lymphocyte-Peyer's
13	642	10.9	1038	2	S06046	integrin alpha-4 c
14	630	10.7	1035	2	I58409	integrin alpha-9 c
15	614.5	10.5	1041	2	T31437	integrin alpha cha
16	579.5	9.9	1054	2	JC7294	alphaP integrin -
17	572.5	9.7	1051	2	A35761	cell surface glyco
18	567.5	9.7	1053	2	I55534	VLA-3 alpha subuni
19	555.5	9.4	1053	2	S44250	integrin alpha-5 c
20	551.5	9.4	1034	2	A36108	integrin alpha-V c
21	539	9.2	1044	2	T10050	integrin alpha-v c
22	533.5	9.1	1072	2	A38457	integrin alpha-6 c
23	532	9.0	1049	2	A27079	fibronectin recept
24	531	9.0	1073	2	B36429	integrin alpha-6 c
25	530	9.0	1048	2	A27421	integrin alpha-5 c
26	529.5	9.0	1051	2	A40021	integrin alpha-3 c
27	524.5	8.9	1091	2	A41543	integrin alpha-6 c
28	514.5	8.8	1044	2	S16516	integrin alpha-8 c
29	497	8.5	1394	2	A29637	position-specific
30	496.5	8.4	1146	2	S40311	integrin - fruit f
31	496	8.4	1039	2	A34269	integrin alpha-2b
32	491	8.4	1137	2	JC5950	integrin alpha-7 c
33	490	8.3	1135	2	I61186	alpha-7 integrin -
34	487.5	8.3	1037	2	A60163	glycoprotein IIb -
35	486	8.3	126	2	B30892	leukocyte adhesion
36	470	8.0	1106	2	S38783	integrin alpha cha
37	466	7.9	1226	2	S44824	F54F2.1 protein -
38	458.5	7.8	1045	2	S60571	integrin alpha v c
39	444.5	7.6	1139	2	S28277	hypothetical prote
40	418	7.1	1115	2	T09403	integrin alpha cha
41	413.5	7.0	1115	2	T09433	integrin alpha cha
42	391	6.7	764	2	I36916	glycoprotein IIb -
43	309	5.3	1086	2	T18523	integrin alpha cha
44	299	5.1	604	2	I36917	glycoprotein IIb -
45	279.5	4.8	272	2	A55348	integrin alpha-1 -

ALIGNMENTS

RESULT 1
RWHU1B
cell surface glycoprotein CD11b precursor [validated] - human
N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Ma
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567
R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.
J. Biol. Chem. 263, 12403-12411, 1988
A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD
B.
A;Reference number: A31108; MUID:88315033; PMID:2457584
A;Accession: A31108
A;Molecule type: mRNA
A;Residues: 1-1153 <COR>
A;Cross-references: UNIPROT:P11215; GB:J03925; NID:g187284; PIDN:AAA59544.1; PID:g307148
A;Note: part of this sequence was confirmed by protein sequencing
R;Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
J. Cell Biol. 106, 2153-2158, 1988
A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor M
A;Reference number: A28915; MUID:88257215; PMID:2454931
A;Accession: A28915
A;Molecule type: mRNA
A;Residues: 1-499,501-965,'P',967-1153 <ARN>
A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AAA594
A;Note: the authors translated the codon TAC for residue 1129 as Thr
A;Note: part of this sequence, including the amino end of the mature protein, was confir
R;Shelley, C.S.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991
A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg
A;Reference number: A41600; MUID:92073318; PMID:1683702
A;Accession: A41600
A;Molecule type: DNA
A;Residues: 1-9 <SHE>
A;Cross-references: GB:M76724; NID:g180018; PIDN:AAA58410.1; PID:g553215
R;Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesi
A;Reference number: A94193; MUID:88190151; PMID:2833753
A;Accession: A30892
A;Molecule type: mRNA
A;Residues: 917-1042 <AR2>
A;Cross-references: GB:M18044
R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
A;Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor
A;Reference number: A32218; MUID:89098893; PMID:2563162
A;Accession: A32218
A;Molecule type: mRNA
A;Residues: 9-1153 <HIC>
A;Cross-references: GB:J04145; NID:g189068; PIDN:AAA59903.1; PID:g386975

A;Note: part of this sequence was confirmed by protein sequencing
R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A;Reference number: A46526; MUID:93123748; PMID:8419480
A;Accession: A46526
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-499,501-1153 <FLE>
A;Cross-references: GB:S52227; NID:g263047; PIDN:AAB24821.1; PID:g263049
A;Note: the last three bases of intron 13, CAG, are included in some but not all mature
A;Note: sequence extracted from NCBI backbone (NCBIP:121963)
R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A;Title: N-terminal sequence of human leukocyte glycoprotein Mo1: conservation across sp
A;Reference number: A90664; MUID:87076671; PMID:3539202
A;Accession: A26091
A;Molecule type: protein
A;Residues: 17-31 <PIE>
A;Experimental source: granulocytes
R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A;Title: Characterization of the myeloid-specific CD11b promoter.
A;Reference number: I52567; MUID:92144986; PMID:1346576
A;Accession: I52567
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-9 <RES>
A;Cross-references: GB:M84477; NID:g180184; PIDN:AAA51960.1; PID:g553219
C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C;Genetics:
A;Gene: GDB:ITGAM; CR3A
A;Cross-references: GDB:120599; OMIM:120980
A;Map position: 16p11.2-16p11.2
A;Note: promoter contains a GATA motif and two Sp1 consensus binding sites
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F;17-1108/Domain: extracellular #status predicted <EXT>
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;465-473/Region: calcium/magnesium binding #status predicted
F;530-538/Region: calcium/magnesium binding #status predicted
F;593-601/Region: calcium/magnesium binding #status predicted
F;1109-1134/Domain: transmembrane #status predicted <TMM>
F;1135-1153/Domain: intracellular #status predicted <INT>
F;86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:||||||||||||||:||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||| |||||||||:| ||||||||||||||:||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 2
S00551
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
N;Alternate names: complement-3 receptor alpha chain
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S00551; I59078
R;Pytela, R.
EMBO J. 7, 1371-1378, 1988
A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the
A;Reference number: S00551; MUID:88312584; PMID:3044779
A;Accession: S00551
A;Molecule type: DNA
A;Residues: 1-1153 <PYT>
A;Cross-references: UNIPROT:P05555; EMBL:X07640; NID:g52982; PIDN:CAA30479.1; PID:g5298
A;Note: the authors translated the codon CAC for residue 569 as Gln

R;Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts,
Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A;Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
A;Reference number: I59078; MUID:86287312; PMID:2942940
A;Accession: I59078
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 11-44 <RES>
A;Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
C;Genetics:
A;Gene: Mac-1
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C;Keywords: cell adhesion; glycoprotein; transmembrane protein
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1106-1129/Domain: transmembrane #status predicted <TMM>

Query Match 75.6%; Score 4447; DB 2; Length 1153;
Best Local Similarity 73.5%; Pred. No. 9.6e-303;
Matches 836; Conservative 147; Mismatches 153; Indels 2; Gaps 2;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         17 FNLDTEHPMTFQENAKGFGQNVVQLGGTSVVVAAPQEAKAVNQTGALYQCDYSTSRCHPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         77 PLQVPPEAVNMSLGLSLAVSTVPQQLLACGPTVHQNCKENTYVNGLCYLFGSNLLRPPQQ 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
Db        137 FPEALRECPQQESDIVFLIDGSGSINNIDFQKMKEFVSTVMEQFKKSKTLFSLMQYSDEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db        197 RIHFTFNDFKRNPSPRSHVSPIKQLNGRTKTASGIRKVVRELFHKTNGARENAAKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db        257 TDGEKFGDPLDYKDVIPEADRAGVIRYVIGVGNAFNKPQSRRELDTIASKPAGEHVFQVD 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        317 NFEALNTIQNQLQEKIFAIEGTQTGSTSSFEHEMSQEGFSASITSNGPLLGSVGSFDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        377 GAFLYTSKDKVTFINTTRVDSDMNDAYLGYASAVILRNRVQSLVLGAPRYQHIGLVVMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        437 ENFGTWEPHTSIKGSQIGSYFGASLCSVDMDADGNTNLILIGAPHYYEKTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        497 PRG-RARWQCEALLHGDQGHPWGRFGAALTVLGDVNGDKLTDVAIGAPGEQENQGAVYIF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        556 YGASIASLSASHSHRIIGAHFSPGLQYFGQSLSGGKDLTMDGLMDLAVGAQGHLLLLRAQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        616 PVLRLEATMEFSPKKVARSVFACQEQVLKNKDAGEVRVCLRVRKNTKDRLREGDIQSTVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        676 YDLALDPVRSRIRAFFDETKNNTRRRTQVFGLMQKCETLKLILPDCVDDSVSPIILRLNY 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        736 TLVGEPLRSFGNLRPVLAMDAQRFFTAMFPFEKNCGNDSICQDDLSITMSAMGLDTLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRL-ACESASSTE 839
Db        796 GPQDFNMSVTLRNDGEDSYGTQVTVYYPSGLSYRKDSASQNPLTKKPWFVKPAESSSSE 855

Qy        840 VSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTE 899
Db        856 GHGALKSTTWNINHPIFPANSEVTFNVTFDVDSHASFGNKLLLKAIVASENNMSRTHKTK 915

Qy        900 FQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        916 FQLELPVKYAIYMIVTSDESSIRYLNFTASEMTSKVIQHQYQFNNLGQRSLPVSVVFWIP 975

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        976 VQINNVTVWDHPQVIFSQNLSSACHTEQKSPPHSNFRDQLERTPVLNCSVAVCKRIQCDL 1035

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1036 PSFNTQEIFNVTLKGNLSFDWYIKTSHGHLLLVSSTEILFNDSAFALLPGQESYVRSKTE 1095

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1096 TKVEPYEVHNPVPLIVGSSIGGLVLLALITAGLYKLGFFKRQYKDMMNEAAPQDAPPQ 1153
```

RESULT 3
RWHU1C
cell surface glycoprotein CD11c precursor - human
N;Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A36584; A35543; S00864
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A;Reference number: A36584
A;Contents: erratum
A;Accession: A36584
A;Molecule type: DNA
A;Residues: 1-1163 <COR>
A;Cross-references: UNIPROT:P20702
A;Note: this revision to the sequence from reference A35543 includes the carboxyl end
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.
A;Reference number: A35543; MUID:90153906; PMID:2303426
A;Accession: A35543
A;Molecule type: DNA
A;Residues: 1-834 <CO2>
A;Note: this sequence has been revised in reference A36584
R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A;Reference number: S00864; MUID:88166645; PMID:3327687
A;Accession: S00864
A;Molecule type: mRNA
A;Residues: 1-755,'L',757-1163 <CO3>
A;Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830
A;Note: part of this sequence was confirmed by protein sequencing
C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my
C;Genetics:
A;Gene: GDB:ITGAX; CD11C
A;Cross-references: GDB:119758; OMIM:151510
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C;Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F;20-1107/Domain: extracellular #status predicted <EXT>
F;149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F;1108-1133/Domain: transmembrane #status predicted <TMM>
F;1134-1163/Domain: intracellular #status predicted <INT>
F;61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match 58.8%; Score 3456; DB 1; Length 1163;
Best Local Similarity 60.6%; Pred. No. 2.5e-233;
Matches 684; Conservative 143; Mismatches 296; Indels 6; Gaps 4;

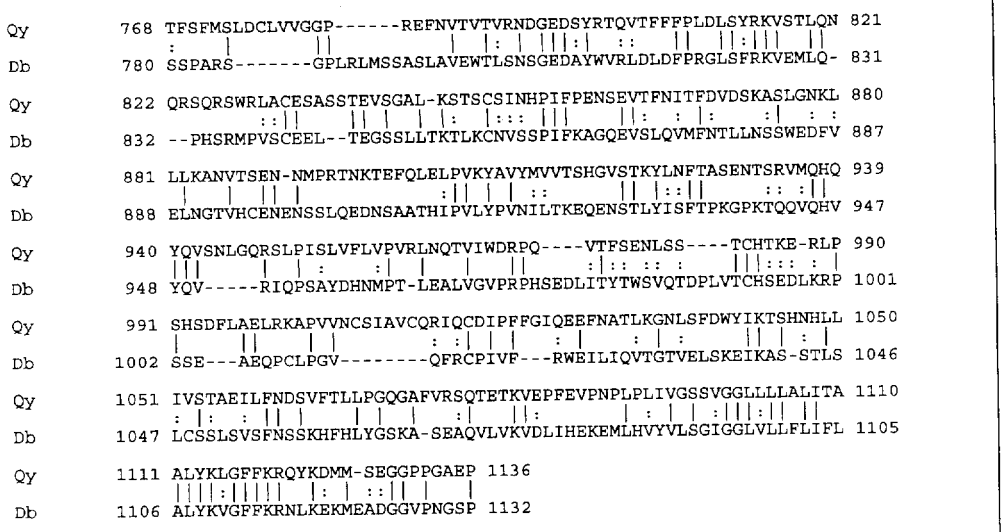
```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAADAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 4
S03308
cell surface glycoprotein CD11a precursor - human
N;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte functior
C;Species: Homo sapiens (man)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: S03308; A47458; A47565; A48759; S36044
R;Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.
J. Cell Biol. 108, 703-712, 1989
A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha subunit
A;Reference number: S03308; MUID:89139587; PMID:2537322
A;Accession: S03308
A;Molecule type: mRNA
A;Residues: 1-1170 <LAR>
A;Cross-references: UNIPROT:P20701; UNIPROT:Q9UBC8; EMBL:Y00796; NID:g31421; PIDN:CAA68
A;Note: part of this sequence was confirmed by protein sequencing
R;Cornwell, R.D.; Gollahon, K.A.; Hickstein, D.D.
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993
A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a) pr
A;Reference number: A47458; MUID:93248261; PMID:8097887
A;Accession: A47458
A;Molecule type: DNA
A;Residues: 1-20 <COR>
A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBIP:130863)
R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993
A;Title: Identification of cell-specific and developmentally regulated nuclear factors
A;Reference number: A47565; MUID:93281759; PMID:8099450
A;Accession: A47565
A;Molecule type: DNA
A;Residues: 1-20 <SHE>
A;Cross-references: GB:M95609
R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.
J. Biol. Chem. 268, 19305-19311, 1993
A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter.
A;Reference number: A48759; MUID:93374910; PMID:8103515
A;Accession: A48759
A;Molecule type: DNA
A;Residues: 1-20 <NUE>
A;Cross-references: EMBL:Z22804; NID:g311405; PIDN:CAA80461.1; PID:g311406
C;Genetics:
A;Gene: GDB:ITGAL; CD11A
A;Cross-references: GDB:119757; OMIM:153370
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen; tr
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted
F;154-317/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 26.1%; Score 1532.5; DB 2; Length 1170;
Best Local Similarity 34.2%; Pred. No. 1.3e-98;
Matches 401; Conservative 209; Mismatches 463; Indels 101; Gaps 37;

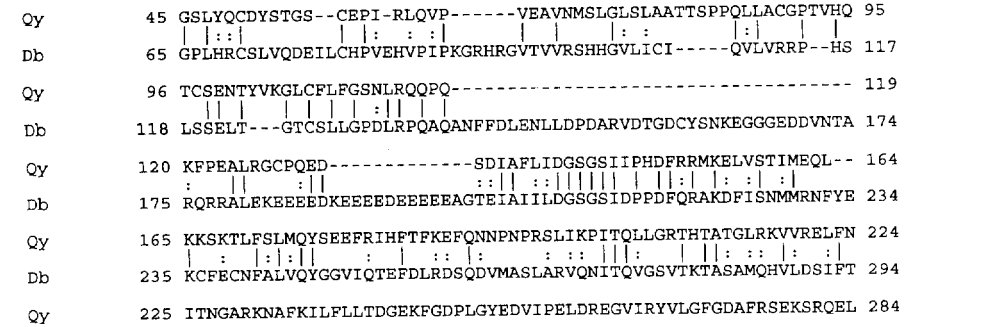
```
Qy          1 FNLDTENAMTFQ--ENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58
Db         26 YNLDVRGARSFSPPRAGRHFGYRVLQV-GNGVIVGAPGE---GNSTGSLYQCQSGTGHCL 81

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLR--- 115
Db         82 PVTLR-GSNYTSKYLGMTLATDPTDGSILACDPGLSRTCDQNTYLSGLCYLFRQNLQGPM 140

Qy        116 -QQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLM 174
```

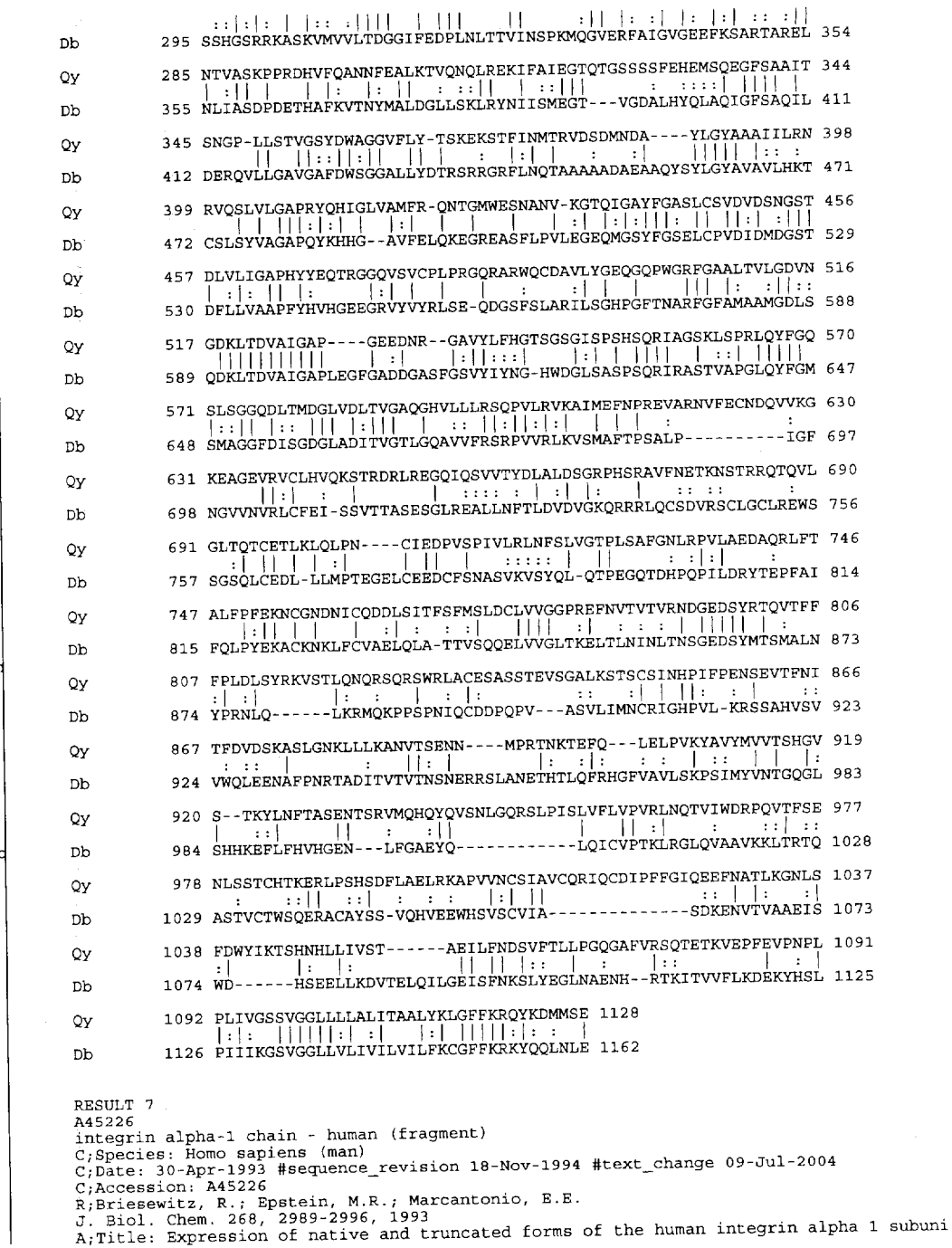
```
Db        141 LQGRPGFQECIKG----NVDLVFLFDGSMSLQPDEFQKILDFMKDVMKKLSNTSYQFAAV 196

Qy        175 QYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAF 234

Db        197 QFSTSYKTEFDFSDYVKWKDPDALLKHVKHMLLLTNTFGAINYVATEVFREELGARPDAT 256

Qy        235 KILFLLTDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRD 294

Db        257 KVLIIITDGE--ATDSGNIDAAKD-----IIRYIIGIGKHFQTKESQETLHKFASKPASE 309

Qy        295 HVFQANNFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVG 354

Db        310 FVKILDTFEKLKDLFTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSRGHAVVGAVG 369

Qy        355 SYDWAGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQH 412

Db        370 AKDWAGGFLDLKADLQDDTFIGNEPLTPEVRAGYLGYTVTWLPSRQKTSLLASGAPRYQH 429

Qy        413 IGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQT 470

Db        430 MGRVLLFQEPQGGGHWSQVQTIHGTQIGSYFGGELCGVDVDQDGETELLLIGAPLFYGEQ 489

Qy        471 RGGQVSVCPLPRGQRARWQCDAV--LYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAP 528

Db        490 RGGRVFIY-----QRRQLGFEEVSELQGDPGYPLGRFGEAITALTDINGDGLVDVAVGAP 544

Qy        529 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTV 588

Db        545 LEE--QGAVYIFNGRHG-GLSPQPSQRIEGTQVLSGIQWFGRSIHGVKDLEGDGLADVAV 601

Qy        589 GAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVV-KGKEAGEVRVCLHVQKSTR 647

Db        602 GAESQMIVLSSRPVVDMVTLMSFSPAEIPVHEVECSYSTSNKMKEGVNITICFQI-KSLY 660

Qy        648 DRLREGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCI 707

Db        661 PQF-QGRLVANLTYTLQLDGHRTRRRGLFPGGRHELRRNIAVT-TSMSCTDFSFHFPVCV 718

Qy        708 EDPVSPIVLRLNFSL---VGTPLS--AFGN-----LRPVLAEDAQRLFTALFPFEKNCGN 757

Db        719 QDLISPINVSLNFSLWEEEGTPRDQRAQGKDIPPILRPSLHSETWEI-----PFEKNCGE 773

Qy        758 DNICQDDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVS 817

Db        774 DKKCEANLRVSFSPARSRALRLTAFASLSVELSLSNLEEDAYWVQLDLHFPPGLSFRKVE 833

Qy        818 TLQNQRSQRSWRLACES--ASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKAS 875

Db        834 ML---KPHSQIPVSCEELPEESRLLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNSS 887

Qy        876 LGNKLLLKANVTSENN----MPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASEN 931

Db        888 WGDSVELHANVTCNNEDSDLLEDNSATTI---IPILYPINILIQDQEDSTLYVSFTPKGP 944

Qy        932 TSRVMQHQYQV---SNLGQRSLP-ISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTK- 986

Db        945 KIHQVKHMYQVRIQPSIHDHNIPTLEAVVGVPQPPSEGPITHQWSVQMEPPV--PCHYED 1002

Qy        987 -ERLPSHSD--FLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIK 1043

Db       1003 LERLPDAAEPCLPGALFRCPVV-----------------FRQEILVQVIGTLELVGEIE 1044

Qy       1044 TSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLL 1103

Db       1045 AS-SMFSLCSSLSISFNSSKHFHLYGSNASL-AQVVMKVDVVYEKQMLYLYVLSGIGGLL 1102

Qy       1104 LLALITAALYKLGFFKRQYKDMMSEG-GPPGAEP 1136

Db       1103 LLLLIFIVLYKVGFFKRNLKEKMEAGRGVPNGIP 1136

RESULT 5
I56126
lymphocyte fuction-associated molecule-1-alpha - mouse
C;Species: Mus musculus (house mouse)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: I56126
R;Kaufmann, Y.; Tseng, E.; Springer, T.A.
J. Immunol. 147, 369-374, 1991
A;Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit
A;Reference number: I56126; MUID:91268576; PMID:2051027
A;Accession: I56126
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1163 <RES>
A;Cross-references: UNIPROT:P24063; GB:M60778; NID:g198785; PIDN:AAA39426.1; PID:g19878
C;Genetics:
A;Gene: LFA-1
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
F;151-315/Domain: von Willebrand factor type A repeat homology <VWA1>

  Query Match             25.8%;  Score 1516.5;  DB 2;  Length 1163;
  Best Local Similarity   33.9%;  Pred. No. 1.7e-97;
  Matches  396;  Conservative  216;  Mismatches  466;  Indels   89;  Gaps   36;

Qy          1 FNLDTENAMTFQENA-RGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEP 59

Db         24 YNLDTRPTQSFLAQAGRHFGYQVLQIEDG-VVVGAPGE---GDNTGGLYHCRTSSEFCQP 79

Qy         60 IRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQ 119

Db         80 VSLH-GSNHTSKYLGMTLATDAAKGSLLACDPGLSRTCDQNTYLSGLCYLFPQSLEGPML 138

Qy        120 KFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEE 179

Db        139 QNRPAYQECMKGKVDLVFLFDGSQSLDRKDFEKILEFMKDVMRKLSNTSYQFAAVQFSTD 198

Qy        180 FRIHFTFKEF-QNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILF 238

Db        199 CRTEFTFLDYVKQNKNPDVLLGSVQPMFLLTNTFRAINYVVAHVFKEESGARPDATKVLV 258

Qy        239 LLTDGEKF--GDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHV 296

Db        259 IITDGEASDKGNISAAHD---------ITRYIIGIGKHFVSVQKQKTLHIFASEPVEEFV 309

Qy        297 FQANNFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSY 356

Db        310 KILDTFEKLKDLFTDLQRRIYAIEGTNRQDLTSFNMELSSSGISADLSKGHAVVGAVGAK 369

Qy        357 DWAGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQHIG 414

Db        370 DWAGGFLDLREDLQGATFVGQEPLTSDVRGGYLGYTVAWMTSRSSRPLLAAGAPRYQHVG 429

Qy        415 LVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRG 472

Db        430 QVLLFQAPEAGGRWNQTQKIEGTQIGSYFGGELCSVDLDQDGEAELLLIGAPLFFGEQRG 489

Qy        473 GQVSVCPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEED 532

Db        490 GRVFTY---QRRQSLFEMVSELQGDPGYPLGRFGAAITALTDINGDRLTDVAVGAPLEE- 545

Qy        533 NRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQG 592

Db        546 -QGAVYIFNGKPG-GLSPQPSQRIQGAQVFPGIRWFGRSIHGVKDLGGDRLADVVVGAEG 603

Qy        593 HVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAG-EVRVCLHVQKSTRDRLR 651

Db        604 RVVVLSSRPVVDVVTELSFSPEEIPVHEVECSYSAREEQKHGVKLKACFRIKPLTPQ--F 661

Qy        652 EGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPV 711

Db        662 QGRLLANLSYTLQLDGHRMRSRGLFPDGSHELSGNTSITP-DKSCLDFHFHFPICIQDLI 720

Qy        712 SPIVLRLNFSLV---GTPLSAFGN-LRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSI 767

Db        721 SPINVSLNFSLLEEEGTPRDQKGRAMQPILRPSIHTV-TKEIPFEKNCGEDKKCEANLTL 779
```



RESULT 6
A53213
integrin alpha-E chain - human
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 09-Jul-2004
C;Accession: A53213
R;Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6016-6025, 1994
A;Title: Molecular cloning of the human mucosal lymphocyte integrin alpha(E) subunit. Un
A;Reference number: A53213; MUID:94164962; PMID:8119947
A;Accession: A53213
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1179 <SHA>
A;Cross-references: UNIPROT:P38570; GB:L25851; NID:g457244; PID:g457245
C;Genetics:
A;Gene: GDB:ITGAE
A;Cross-references: GDB:330801
A;Map position: 17p13
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
F;199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 19.2%; Score 1128; DB 2; Length 1179;
Best Local Similarity 28.7%; Pred. No. 2.9e-70;
Matches 338; Conservative 213; Mismatches 454; Indels 172; Gaps 39;





RESULT 7
A45226
integrin alpha-1 chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A45226
R;Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.
J. Biol. Chem. 268, 2989-2996, 1993
A;Title: Expression of native and truncated forms of the human integrin alpha 1 subuni

A;Reference number: A45226; MUID:93155124; PMID:8428973
A;Accession: A45226
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1151 <BRI>
A;Cross-references: UNIPROT:P56199
A;Experimental source: hepatoblastoma cell line HepG2
A;Note: sequence extracted from NCBI backbone (NCBIP:124326)
F;142-317/Domain: von Willebrand factor type A repeat homology <VWA1>

Query Match 18.5%; Score 1085.5; DB 2; Length 1151;
Best Local Similarity 27.6%; Pred. No. 2.6e-67;
Matches 342; Conservative 209; Mismatches 488; Indels 201; Gaps 44;

```
Qy          1 FNLDTENAMTFQENARG-FGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQCDYSTGS 56
              ||:| :|:||| || :| | :| |::|:|      |: | :|:|    |
Db          1 FNVDVKNSMTFSGPVEDMFGYTVQQYENEEGKWVLIGSPLVGQPKNRTGDVYKCPVGRGE 60

Qy         57 CEP-IRLQVPVEA---------VNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGL 106
               | ::| :||          ||: | :|  |      |||||     |    | |:
Db         61 SLPCVKLDLPVNTSIPNVTEVEKENMTFGSTL-VTNPNGGFLACGPLYAYRCGHLHYTTGI 119

Qy        107 CFLFGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKK 166
              |       :      |  :: |  :  ||  ::||| || | |      : : : : ||:
Db        120 CSDVSPTFQVVNSIAP--VQECSTQ-LDIVIVLDGSNSIYPWD-----SVTAFLNDLLKR 171

Qy        167 -----SKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLGR-THTATGLRKVVR 220
                  :|   ::|| |     |   :: :       | | |  || | || |     :
Db        172 MDIGPKQTQVGIVQYGENVTHEFNLNKYSSTEEVLVAAKKIVQRGGRQTMTALGTDTARK 231

Qy        221 ELFNITNGARKNAFKILFLLTDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFR---- 276
              | |    |||:   |:: ::|||| |   : || : : | : |:  ::
Db        232 EAFTEARGARRGVKKVMVIVTDGESH-DNHRLKKVIQDCEDENIQRFSIAILGSYNRGNL 290

Qy        277 -SEKSRQELNTVASKPPRDHVFQANNFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMS 335
               :||  :|: ::||:|   | |  ::  || |:    | |:|||:| |   |::||| |||
Db        291 STEKFVEEIKSIASEPTEKHFFNVSDELALVTIVKTLGERIFALEATADQSAASFEMEMS 350

Qy        336 QEGFSAAITSNGPLLSTVGSYDWAGGVFLYTSKE-----KSTF-INMTRVDSDMNDAYLG 389
              | ||||  : :  :|  ||:||| | | :  :  :     : | ||||:|||: | :
Db        351 QTGFSAHYSQDWVMLGAVGAYDWNGTVVMQKASQIIIPRNTTFNVESTKKNEPL-ASYLG 409

Qy        390 YAAAIILRNRVQSL-VLGAPRYQHIGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSV 448
              |       :  | :| ||| | | |  ::|    |  :    : | |||:|||: | :
Db        410 YTVNSATASSGDVLYIAGQPRYNHTGQVIIYRMEDGNIKILQTLSGEQIGSYFGSILTTT 469

Qy        449 DVDSNGSTDLVLIGAPHY---------------YEQTR-GGQVSVCPLPRGQRARWQCDA 492
              |:| : :||::|:||| |                |||   |:|: |: :   :  | ::
Db        470 DIDKDSNTDILLVGAPMYMGTEKEEQGKVYVYALNQTRFEYQMSLEPIKQTCCSSRQHNS 529

Qy        493 VLYGEQGQPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPS 551
               : :| | ||| |:  : |:| |    |: |||| |:|: ||||::|| ||  |
Db        530 CTTENKNEPCGARFGTAIAAVKDLNLDGFNDIVIGAPLEDDHGGAVYIYHG-SGKTIRKE 588

Qy        552 HSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIMEF 611
              ::|||         |::|||: |  ||  ||| |:|:|  |    |  |  |:|  ||  | |
Db        589 YAQRIPSGGDGKTLKFFGQSIHGEMDLNGDGLTDVTIGGLGGAALFWSRDVAVVKVTMNF 648

Qy        612 NPREVARNVFECNDQVVKGKEAG--EVRVCLHVQ-KSTRDRLREGQIQSVVTYDLALDSG 668
              | :|      |:   ::|||       ||  |: ||  | : |  :|     | : |||
Db        649 EPNKVNIQKKNCH---MEGKETVCINATVCFEVKLKSKEDTIYEADLQ----YRVTLDSL 701

Qy        669 RPHSRAVFNET------KNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNFSL 722
              |  ||: |: |       :| | |:::     :  :  :  : :| |    : |:|:|
Db        702 RQISRSFFSGTQERKVQRNITVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDFNL 753

Qy        723 VGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVGGP 782
              |  :      |||  :              || |:|||    |  |||:  :     | |:|
Db        754 T-DPENG-----PVLDDSLPNSVHEYIPFAKDCGNKEKCISDLSLHVATTEKDLLIVRSQ 807

Qy        783 RE-FNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEVS 841
              : |||::||:|  : :| |:    :  :| :  :  :|        :|||
Db        808 NDKFNVSLTVKNTKDSAYNTRTIVHYSPNLVFSGIEAIQKD--------SCESN------ 853
```

Qy 842 GALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGN-KLLLKANVTSENNMPRTNKTEF 900
: :| : :| ||| | | :: : | | || :
Db 854 ---HNITCKVGYPFLRRGEMVTFKILFQFNTSYLMENVTIYLSATSDSEEPPETLSDNVV 910

Qy 901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSN------LGQRS----- 949
: :|||| | : | | ::: |:| |: : | |
Db 911 NISIPVKYEVGLQFYS-SASEYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSFPM 969

Qy 950 --LPISLVF----------LVPVRLNQTVIWDRPQVTFSENLSSTCHTKE---------R 988
| :|: | | | |: ||| : | | :
Db 970 PELKLSISFPNMTSNGYPVLYPTGLSS----------SENANCRPHIFEDPFSINSGKK 1018

Qy 989 LPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQE----------EFNA 1030
: : :| |:: :::|: | |:: | : |
Db 1019 MTTSTD---HLKRGTILDCNTCKFATITCNLTSSDISQVNVSLILWKPTFIKSYFSSLNL 1075

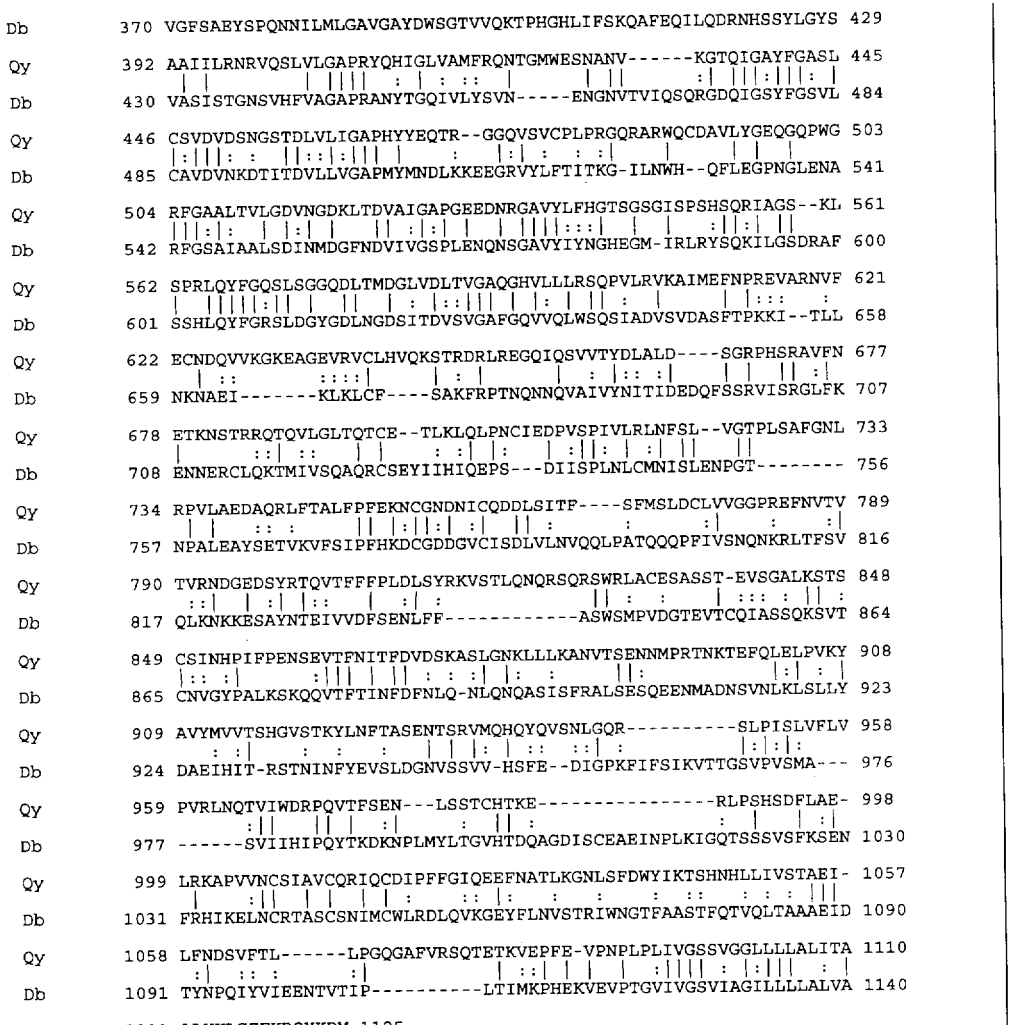
Qy 1031 TLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPPEVPNP 1090
|::| | | | |::|::: : : :: :|
Db 1076 TIRGEL-------RSENASLVLSSSN----------QKRELAIQISKDGLPGR 1111

Qy 1091 LPL--IVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
:|| |: |: ||||| |: ||:|:||||| | | :
Db 1112 VPLWVILLSAFAGLLLLMLLILALWKIGFFKRPLKKKMEK 1151

RESULT 8
I45914
integrin alpha 2 subunit - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C;Accession: I45914
R;Kamata, T.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A;Title: Identification of putative ligand binding sites within the I-domain of integrin
A;Reference number: A54402; MUID:94193647; PMID:7511592
A;Accession: I45914
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1170 <KAM>
A;Cross-references: UNIPROT:P53710; GB:L25886; NID:g439695; PIDN:AAB59255.1; PID:g439696
F;161-336/Domain: von Willebrand factor type A repeat homology <VWA2>

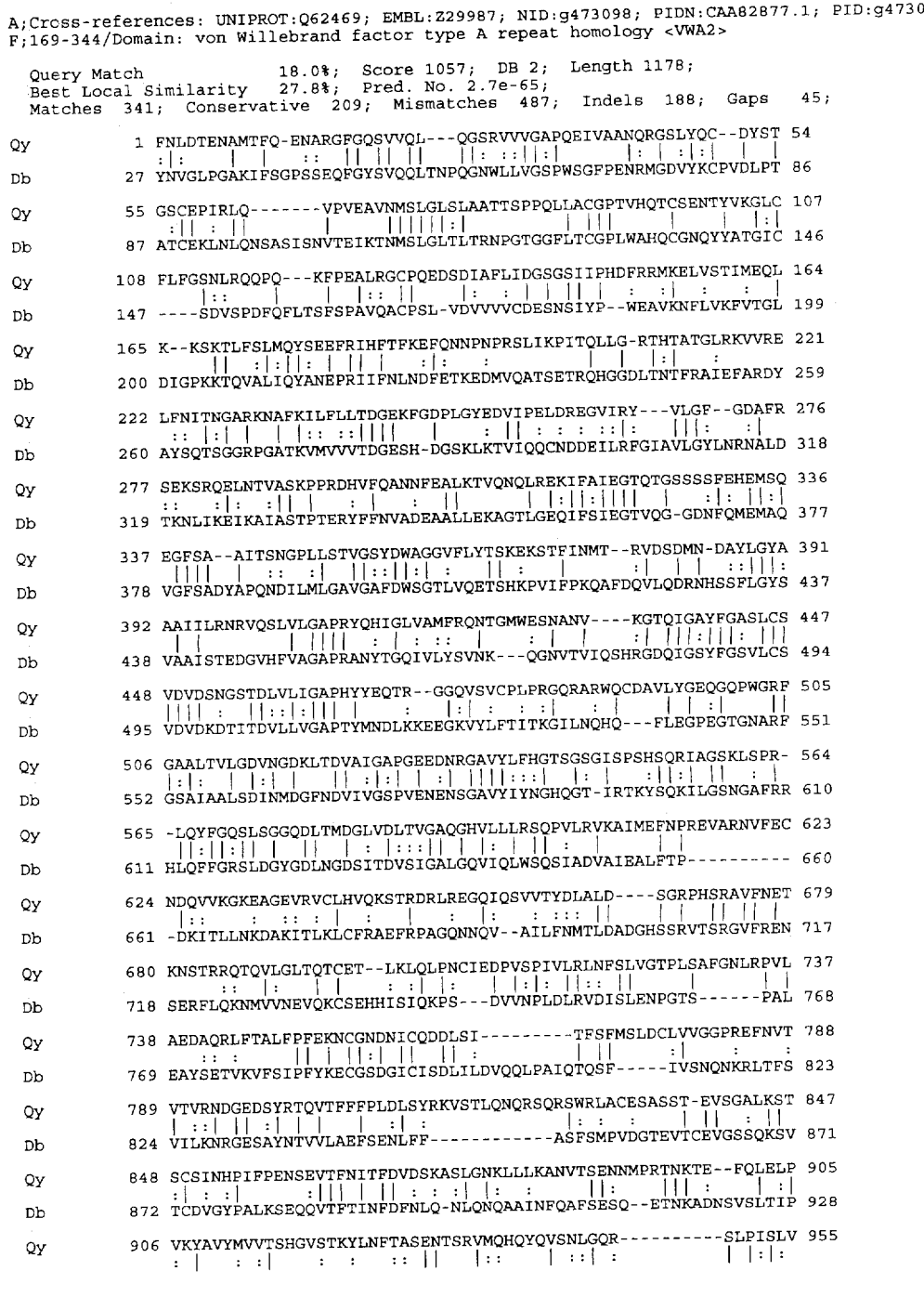
Query Match 18.2%; Score 1071; DB 2; Length 1170;
Best Local Similarity 27.5%; Pred. No. 2.8e-66;
Matches 334; Conservative 218; Mismatches 495; Indels 168; Gaps 48;

```
Qy          1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQC--DYST 54
              :|:    |  |   ::  || :| |    :|: ::||:|      |: | :|:|  | ||
Db         19 YNVGLPKAKIFSGPSSEQFGYAVQQFINPKGNWLLVGSPWSGFPKNRMGDVYKCPVDLST 78

Qy         55 GSCEPIRLQ-------VPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLC 107
              :|| : ||        |      ||||||:|          | |||   | |    |  |:|
Db         79 TTCEKLNLQTSTSMSNVTEMKTNMSLGLTLTRNVGTGGFLTCGPLWAQQCGSQYYTTGVC 138

Qy        108 FLFGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLK-- 165
              :  |     |  |:: ||    |:  : | | || | |   :|  :  :: |
Db        139 SDVSPDF-QLRTSFAPAVQTCP-SFIDVVVVCDESNSIYPWD--AVKNFLEKFVQGLDIG 194

Qy        166 KSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLIKPITQLL----GRTHTATGLRKVVRE 221
              :||   |:||:   |: |    |::          :||  :|       |:|   ::
Db        195 PTKTQMGLIQYANNPRVVFNLNTFKSKD---EMIKATSQTFQYGGDLTNTFKAIQYARDT 251

Qy        222 LFNITNGARKNAFKILFLLTDGEKFGDPLGYEDVIPELDREGVIRY---VLGF--GDAFR 276
              ::   | |  | |:: ::||||    |     : || : ::: ::|:   |||:   :|
Db        252 AYSTAAGGRPGATKVMVVVTDGESH-DGSKLKAVIDQCNKDNILRFGIAVLGYLNRNALD 310

Qy        277 SEKSRQELNTVASKPPRDHVFQANNFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQ 336
              ::   :|:  :||| |  | |  ::      |        : |:||:|||| |   :|: ||||
Db        311 TKNLIKEIKAIASIPTERHFFNVSDEADLLEKAGTIGEQIFSIEGTVQG-GDNFQMEMSQ 369

Qy        337 EGFSAAIT--SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYA 391
              ||||  :  :|  :|  ||:|||:| |   |       |      ::  | |  :||||:
```



RESULT 9
S44142
VLA-2 protein homolog - mouse
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S44142
R;Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze
submitted to the EMBL Data Library, January 1994
A;Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not
A;Reference number: S44142
A;Accession: S44142
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1178 <EDE>
A;Cross-references: UNIPROT:Q62469; EMBL:Z29987; NID:g473098; PIDN:CAA82877.1; PID:g4730
F;169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.0%; Score 1057; DB 2; Length 1178;
Best Local Similarity 27.8%; Pred. No. 2.7e-65;
Matches 341; Conservative 209; Mismatches 487; Indels 188; Gaps 45;



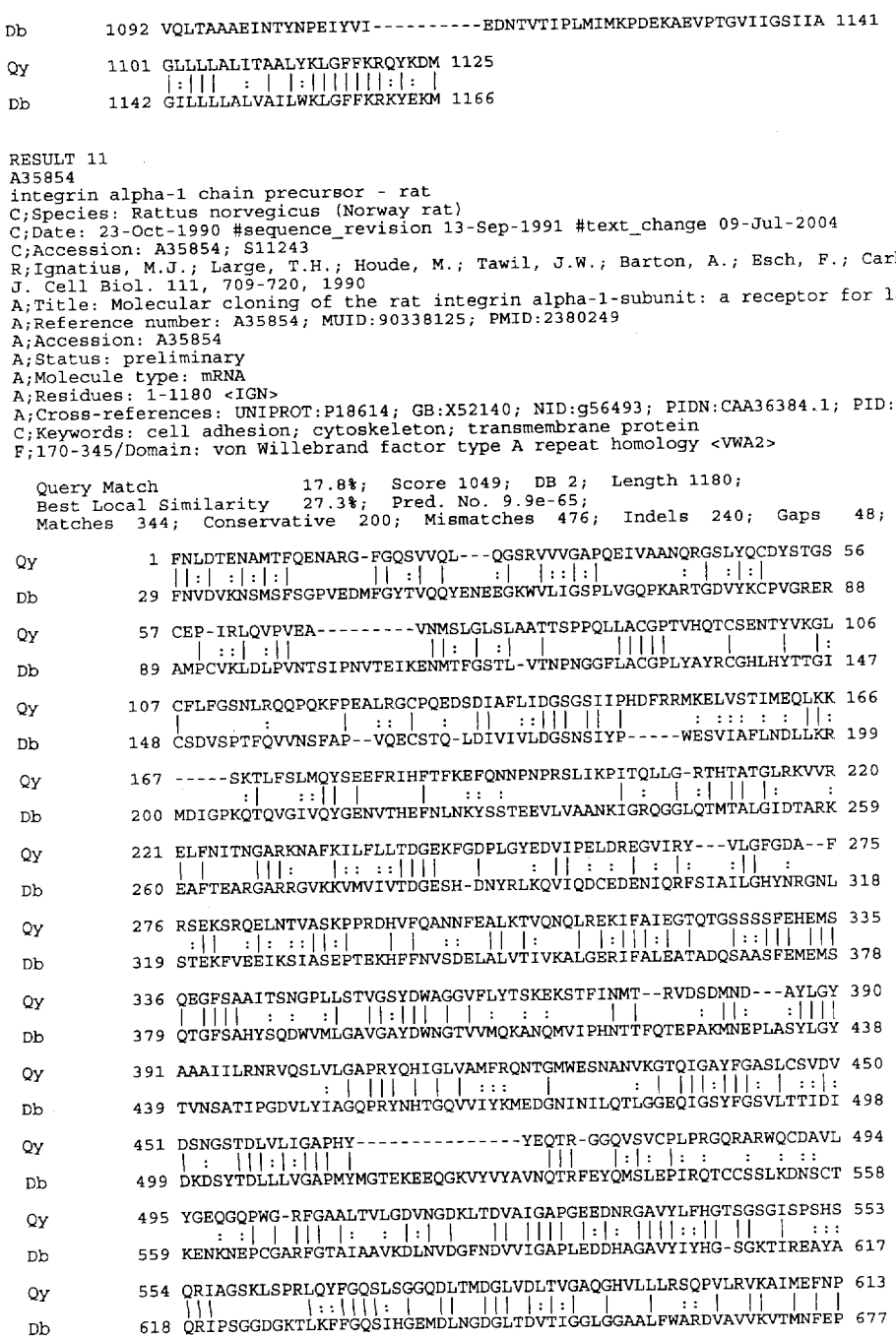
```
Db        929 LLYDAELHLT-RSTNINFYEISSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVSMA 984

Qy        956 FLV----------------------------------PVRLNQTVIWDRPQVTF-SENLS 980
              :                                  |::|  |     | |:| :||
Db        985 LVTIHIPQYTKEKNPLLYLTGIQTDQAGDISCTAEINPLKLPHTA----PSVSFKNENFR 1040

Qy        981 STCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDW 1040
              ||||                        ::|    |  | | :    :: |:   :   :
Db       1041 ---HTKE----------------LDCRTTSCSNITCWLKDLHMKAEYFINVTTRVWNRT 1080

Qy       1041 YIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVG 1100
              :  ::   : : :|||  ::    ::       :      |  |||    :|:|| :
Db       1081 FAASTFQTVQLTAAAEIDTHNPQLFVIEENAVTIPLMIMKPTEKAEVPT--GVIIGSIIA 1138

Qy       1101 GLLLLALITAALYKLGFFKRQYKDM 1125
              |:|||  :|| |:|||||||||| |
Db       1139 GILLLLAMTAGLWKLGFFKRQYKKM 1163
```

RESULT 10
A33998
integrin alpha-2 chain precursor - human
N;Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 09-Jul-2004
C;Accession: A33998; B56793; A53117
R;Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet
A;Reference number: A33998; MUID:89308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: UNIPROT:P17301; GB:X17033; NID:g33906; PIDN:CAA34894.1; PID:g33907
A;Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R;Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIc, GPIc*, GPIIa an
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16,'V',18-21 <ZUT>
A;Cross-references: GB:L24121; NID:g400342; PIDN:AAA16619.2; PID:g4583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TMM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

  Query Match   17.9%; Score 1054; DB 2; Length 1181;
  Best Local Similarity   26.8%; Pred. No. 4.4e-65;
  Matches 328; Conservative 215; Mismatches 494; Indels 188; Gaps   44;

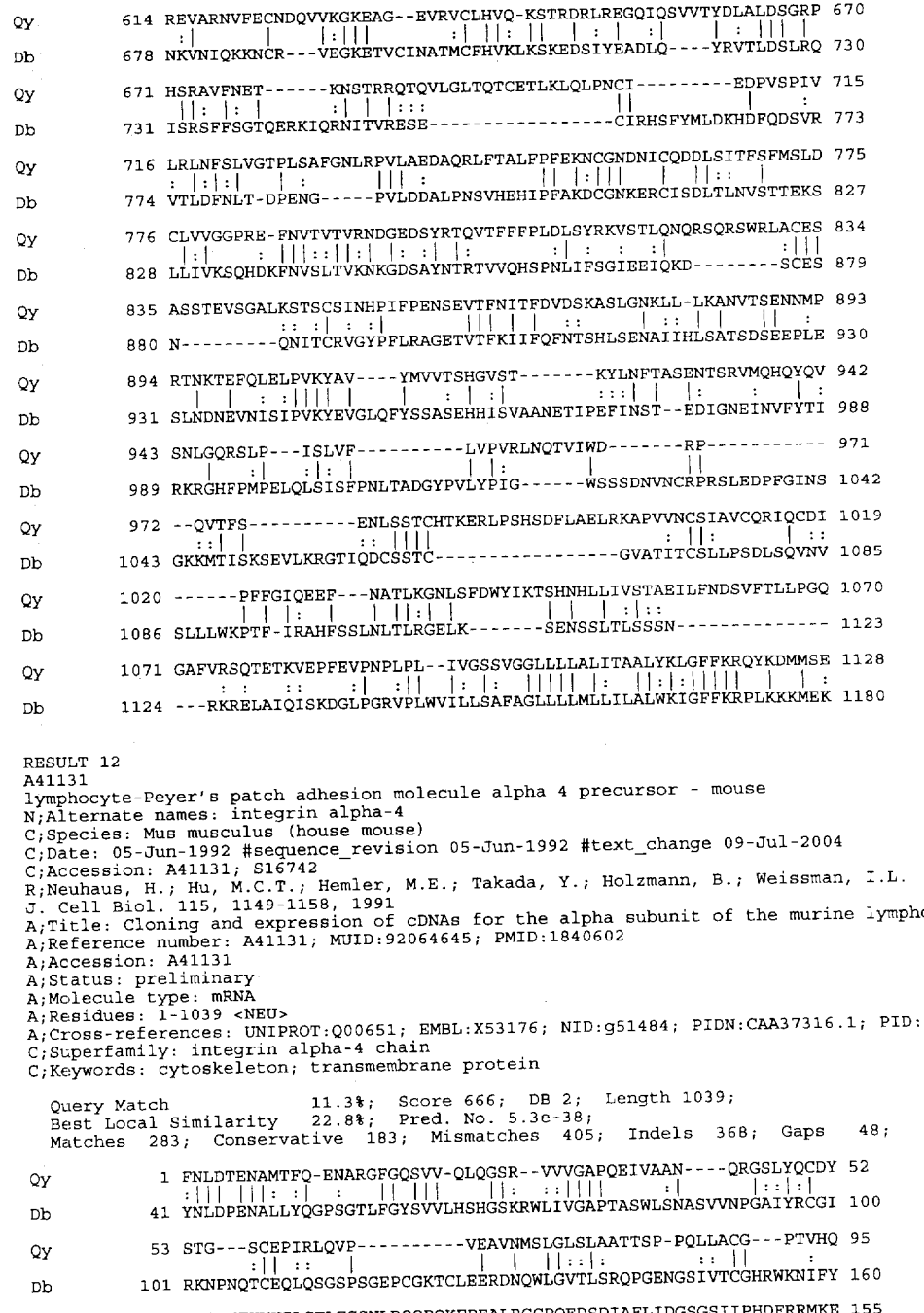
```
Qy          1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQC--DYST 54
              :|:    |  |   ::  || :| |    :|: ::||:|      |: | :|:|  | ||
Db         30 YNVGLPEAKIFSGPSSEQFGYAVQQFINPKGNWLLVGSPWSGFPENRMGDVYKCPVDLST 89

Qy         55 GSCEPIRLQ-------VPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLC 107
```

```
              :|| : ||        |     |||||| |          | |||   | |     |  |:|
Db         90 ATCEKLNLQTSTSIPNVTEMKTNMSLGLILTRNMGTGGFLTCGPLWAQQCGNQYYTTGVC 149

Qy        108 FLFGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLK-- 165
              :  |    |  | : ||    |:  : | | ||| | |   :|  :   :: |
Db        150 SDISPDF-QLSASPSPATQPCPSL-IDVVVVCDESNSIYPWD--AVKNFLEKFVQGLDIG 205

Qy        166 KSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLG-RTHTATGLRKVVRELFN 224
              :||   |:||:   |: |    ::              :|  |  |:|   ::   :  ::
Db        206 PTKTQVGLIQYANNPRVVFNLNTYKTKEEMIVATSQTSQYGGDLTNTFGAIQYARKYAYS 265

Qy        225 ITNGARKNAFKILFLLTDGEKFGDPLGYEDVIPELDREGVIRY---VLGF--GDAFRSEK 279
              :| |::| |:: ::|||| |   : || : : : ::|:   |||:  :|  ::
Db        266 AASGGRRSATKVMVVVTDGESH-DGSMLKAVIDQCNHDNILRFGIAVLGYLNRNALDTKN 324

Qy        280 SRQELNTVASKPPRDHVFQANNFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGF 339
              :|:  :|| |  : |  ::  ||       | |:||:|||| |   :|: |||| ||
Db        325 LIKEIKAIASIPTERYFFNVSDEAALLEKAGTLGEQIFSIEGTVQG-GDNFQMEMSQVGF 383

Qy        340 SAAITSNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYAAAI 394
              ||  :|     :|  ||:: |:| :  ||       |       ::  | |  :|||: |
Db        384 SADYSSQNDILMLGAVGAFGWSGTIVQKTSHGHLIFPKQAFDQILQDRNHSSYLGYSVAA 443

Qy        395 ILRNRVQSLVLGAPRYQHIGLVAMFRQNTGMWESNANV------KGTQIGAYFGASLCSV 448
              |       | |||| : | : ::  |      | |:      :| |||:|||: ||||
Db        444 ISTGESTHFVAGAPRANYTGQIVLYSVN-----ENGNITVIQAHRGDQIGSYFGSVLCSV 498

Qy        449 DVDSNGSTDLVLIGAPHYYEQTR--GGQVSVCPLPRGQRARWQCDAVLYGEQGQPWGRFG 506
              ||| :  ||::|:|||  |    :  |:|  :  : :|   : |    | | :|    |||
Db        499 DVDKDTITDVLLVGAPMYMSDLKKEEGRVYLFTIKKGILGQHQ---FLEGPEGIENTRFG 555

Qy        507 AALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGS--KLSPR 564
              :|:  | |:| |   || :|:| | :|||:::|  |: |   :||:| ||
Db        556 SAIAALSDINMDGFNDVIVGSPLENQNSGAVYIYNGHQGT-IRTKYSQKILGSDGAFRSH 614

Qy        565 LQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECN 624
              |||||:|| |  ||  | : |:::|| | |: | || :  |      | |  ::   :  |
Db        615 LQYFGRSLDGYGDLNGDSITDVSIGAFGQVVQLWSQSIADVAIEASFTPEKI--TLVNKN 672

Qy        625 DQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVTYDLALD----SGRPHSRAVFNETK 680
              |::       :::|     | : |  :   |  : |:: ||    | |  || :| |
Db        673 AQII-------LKLCF----SAKFRPTKQNNQVAIVYNITLDADGFSSRVTSRGLFKENN 721

Qy        681 NSTRRQTQVLGLTQTC--ETLKLQLPNCIEDPVSPIVLRLNFSLVGTPLSAFGNLRPVLA 738
              ::  |:  |:|    : :| |:   | |: : ||:: ||       |       | |
Db        722 ERCLQKNMVVNQAQSCPEHIIYIQEPS---DVVNSLDLRVDISLENPGTS------PALE 772

Qy        739 EDAQRLFTALFPFEKNCGNDNICQDDLSITF----SFMSLDCLVVGGPREFNVTVTVRND 794
              ::       || |:|| | :|  || :       :     :|    :    :||::|
Db        773 AYSETAKVFSIPFHKDCGEDGLCISDLVLDVRQIPAAQEQPFIVSNQNKRLTFSVTLKNK 832

Qy        795 GEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASST-EVSGALKSTSCSINH 853
              | :| | :   |  :| :              |: |  :    | :|: : || :| : :
Db        833 RESAYNTGIVVDFSENLFF-----------ASFSLPVDGTEVTCQVAASQKSVACDVGY 880

Qy        854 PIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEFQLELPVKYAVYMV 913
              |      :||| | || : : :| |:  |     ||:          |:|: |   :
Db        881 PALKREQQVTFTINFDFNLQ-NLQNQASLSFQALSESQEENKADNLVNLKIPLLYDAEI- 938

Qy        914 VTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQR----------SLPISLVFLV----- 958
              |   : :||    :  |   :  ::| :           |:|:|:  ::
Db        939 ---HLTRSTNINFYEISSDGNVPSIVHSFEDVGPKFIFSLKVTTGSVPVSMATVIIHIPQ 995

Qy        959 ----------------------------PVRLNQTVIWDRPQVTF-SENLSSTCHTKER 988
                                         |::: ||       |:| |||     ||||
Db        996 YTKEKNPLMYLTGVQTDKAGDISCNADINPLKIGQT----SSSVSFKSENFR---HTKE- 1047

Qy        989 LPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNH 1048
                                   :||  | |  :  | :   :: |:    :   :  ::
Db       1048 ---------------LNCRTASCSNVTCWLKDVHMKGEYFVNVTTRIWNGTFASSTFQT 1091

Qy       1049 LLIVSTAEI-LFNDSVFTLLPGQGAFVRSQTETKVEPPEVPNP-----LP--LIVGSSVG 1100
              : : : |||  :|  :: :          :  |   |  :  |     :|  :|:|| :
```



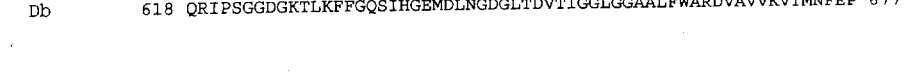
RESULT 11
A35854
integrin alpha-1 chain precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 09-Jul-2004
C;Accession: A35854; S11243
R;Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto,
J. Cell Biol. 111, 709-720, 1990
A;Title: Molecular cloning of the rat integrin alpha-1-subunit: a receptor for laminin a
A;Reference number: A35854; MUID:90338125; PMID:2380249
A;Accession: A35854
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1180 <IGN>
A;Cross-references: UNIPROT:P18614; GB:X52140; NID:g56493; PIDN:CAA36384.1; PID:g56494
C;Keywords: cell adhesion; cytoskeleton; transmembrane protein
F;170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 17.8%; Score 1049; DB 2; Length 1180;
Best Local Similarity 27.3%; Pred. No. 9.9e-65;
Matches 344; Conservative 200; Mismatches 476; Indels 240; Gaps 48;



RESULT 12
A41131
lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
N;Alternate names: integrin alpha-4
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004
C;Accession: A41131; S16742
R;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.L.
J. Cell Biol. 115, 1149-1158, 1991
A;Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte
A;Reference number: A41131; MUID:92064645; PMID:1840602
A;Accession: A41131
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1039 <NEU>
A;Cross-references: UNIPROT:Q00651; EMBL:X53176; NID:g51484; PIDN:CAA37316.1; PID:g5148
C;Superfamily: integrin alpha-4 chain
C;Keywords: cytoskeleton; transmembrane protein

Query Match 11.3%; Score 666; DB 2; Length 1039;
Best Local Similarity 22.8%; Pred. No. 5.3e-38;
Matches 283; Conservative 183; Mismatches 405; Indels 368; Gaps 48;

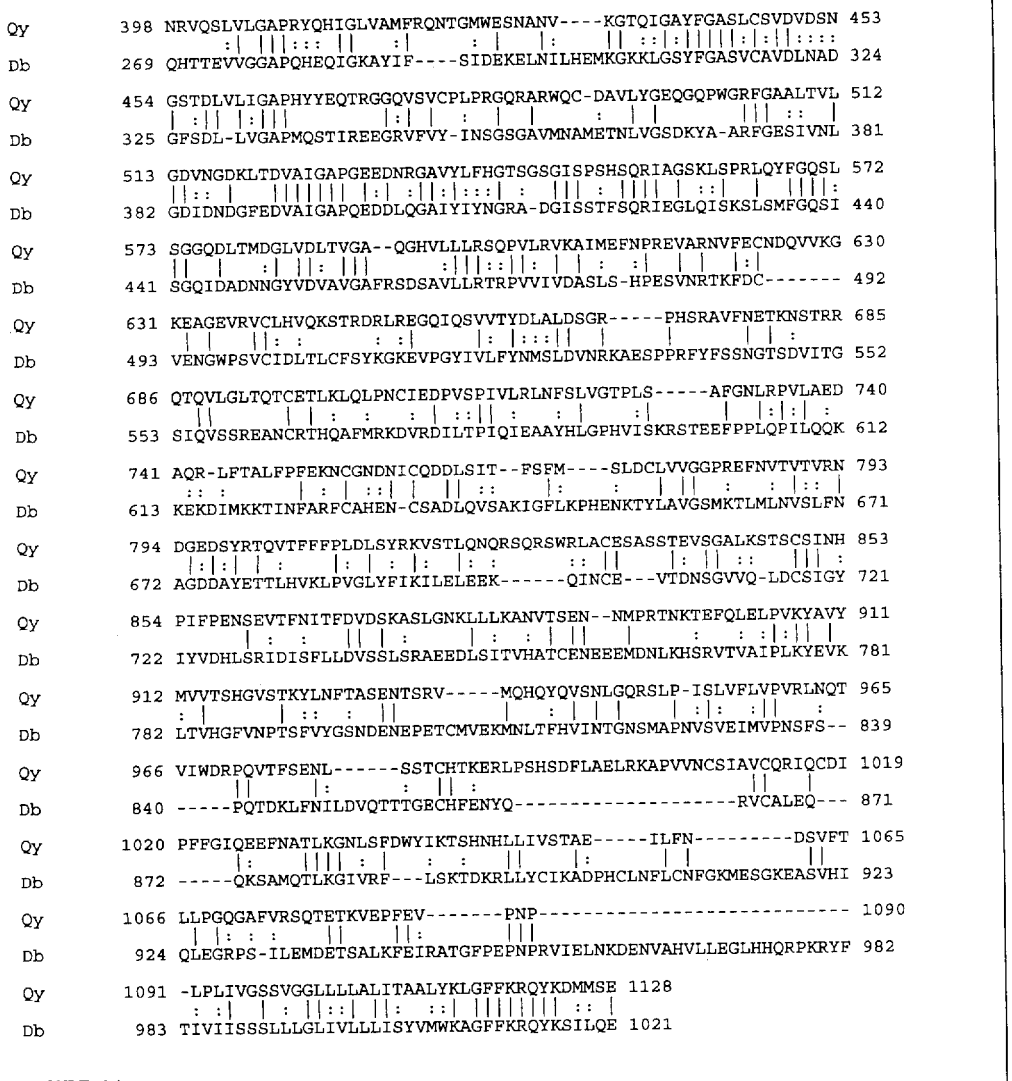


```
Db        161 MKSDNKLPTGICYVMPSDLRTELSK----------------------------RMAP 189

Qy        156 LVSTIMEQLKKSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGL 215

Db        190 C------------------------------------------------------------ 190

Qy        216 RKVVRELFNITNGARKNAFKILFLLTDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAF 275

Db        191 ----------------------------------YKD-----------YTRKFGENF 202

Qy        276 RSEKSRQELNTVASKPPRDHVFQANNFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMS 335

Db        203 AS------------------------------------------------------------C 205

Qy        336 QEGFSAAITSNGPLLSTVGSYDWAGGVFLY--TSKEKSTFINMTRVDSDMNDAYLGYA-- 391

Db        206 QAGISSFYTQDLIVMGAPGSSYWTGTVFVYNITTNQYKAFVD--RQNQVKFGSYLGYSVG 263

Qy        392 AAIILRNRVQSLVLGAPRYQHIGLVAMFRQNTGMWESNANV----KGTQIGAYFGASLCS 447

Db        264 AGHFRSPHTTEVVGGAPQHEQIGKAYIF----SIDENELNIVYEMKGKKLGSYFGASVCA 319

Qy        448 VDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPLPRGQRA-RWQCDAVLYGEQGQPWGRFG 506

Db        320 VDLNADGFSDL-LVGAPMQSTIREEGRVFVY-INSGMGAVMVEMERVLVGSDKYA-ARFG 376

Qy        507 AALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQ 566

Db        377 ESIANLGDIDNDGFEDIAIGAPQEDDLRGAVYIYNGRV-DGISSTYSQRIEGQQISKSLR 435

Qy        567 YFGQSLSGGQDLTMDGLVDLTVGA--QGHVLLLRSQPVLRVKAIMEFNPREVARNVFECN 624

Db        436 MFGQSISGQIDADNNGYVDVAVGAFQSDSAVLLRTRPVVIVEASLS-HPESVNRTKFDCT 494

Qy        625 DQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVTYDLALDSGR---PHSRAVF--NET 679

Db        495 -------ENGLPSVCMHLTLCFSYKGKEVPGYIVLFYNVSLDVHRKAESPSRFYFFSNGT 547

Qy        680 KNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNFSL-----VGTPLSAFGNLR 734

Db        548 SDVITGSIRVSSSGEKCRTHQAFMRKDVRDILTPIHVEATYHLGHHVITKRNTEEFPPLQ 607

Qy        735 PVLAEDAQR-LFTALFPFEKNCGNDNICQDDLSITFS------FMSLDCLVVGGPREFNV 787

Db        608 PILQQKKEKDVIRKMINFARFCAYEN-CSADLQVSAKVGFLKPYENKTYLAVGSMKTIML 666

Qy        788 TVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEVSGALKST 847

Db        667 NVSLFNAGDDAYETTLNVQLPTGLYFIKILDLEEK------QINCE---VTESSGIVK-L 716

Qy        848 SCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSEN--NMPRTNKTEFQLELP 905

Db        717 ACSLGYIYVDRLSRIDISFLLDVSSLSRAHEDLSISVHASCENEGELDQVRDNRVTLTIP 776

Qy        906 VKYAVYMVVTSHGV--STKYLNFTASENTSRVMQHQ-----YQVSNLGQRSLP-ISLVFL 957

Db        777 LRYEV--MLTVHGLVNPTSFVYGSSEENEPETCMAEKLNLTFHVINTGISMAPNVSVKIM 834

Qy        958 VPVRLNQTVIWDRP--QVTFSENLSSTCHTKE---------------RLPSHSDFLAELR 1000

Db        835 VP---NSFLPQDDKLFNVLDVQTTTGQCHFKHYGRECTFAQQKGIAGTLTDIVKFLSKTD 891

Qy       1001 KAPVVNCSIA--VCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEIL 1058

Db        892 KR-LLYCMKADQHCLDFLCN---FGKMESG-------------KEASVHIQLEGRPSIL 933

Qy       1059 FNDSVFTLLPGQGAFVRSQTETKVEPFEVPNP---------------------------- 1090

Db        934 EMDETSSL----------KFEIKATAFPEPHPKVIELNKDENVAHVFLEGLHHQRPKRHF 983

Qy       1091 -LPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128

Db        984 TIIIITISLLLGLIVLLLISCVMWKAGFFKRQYKSILQE 1022
```

RESULT 13
S06046
integrin alpha-4 chain precursor - human
N;Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain
C;Species: Homo sapiens (man)
C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C;Accession: S06046; A39355; D28018
R;Takada, Y.; Elices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989
A;Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other inte
A;Reference number: S06046; MUID:89356603; PMID:2788572
A;Accession: S06046
A;Molecule type: mRNA
A;Residues: 1-1038 <TAK>
A;Cross-references: UNIPROT:P13612; GB:X16983; EMBL:X15356; NID:g33945; PIDN:CAA34852.1
R;Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A;Title: Characterization of the alpha-4 integrin gene promoter.
A;Reference number: A39355; MUID:91239513; PMID:2034655
A;Accession: A39355
A;Molecule type: DNA
A;Residues: 1-93 <ROS>
A;Cross-references: GB:M62841
R;Takada, Y.; Strominger, J.L.; Hemler, M.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987
A;Title: The very late antigen family of heterodimers is part of a superfamily of molecu
A;Reference number: A94151; MUID:87204112; PMID:3033641
A;Accession: D28018
A;Molecule type: protein
A;Residues: 40-50,'E',52-53 <TA2>
C;Genetics:
A;Gene: GDB:ITGA4; CD49D
A;Cross-references: GDB:128032; OMIM:192975
A;Map position: 2q31-2q32
C;Superfamily: integrin alpha-4 chain
C;Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transme
F;1-39/Domain: signal sequence #status predicted <SIG>
F;40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.9%; Score 642; DB 2; Length 1038;
Best Local Similarity 23.6%; Pred. No. 2.5e-36;
Matches 293; Conservative 170; Mismatches 408; Indels 368; Gaps 51;

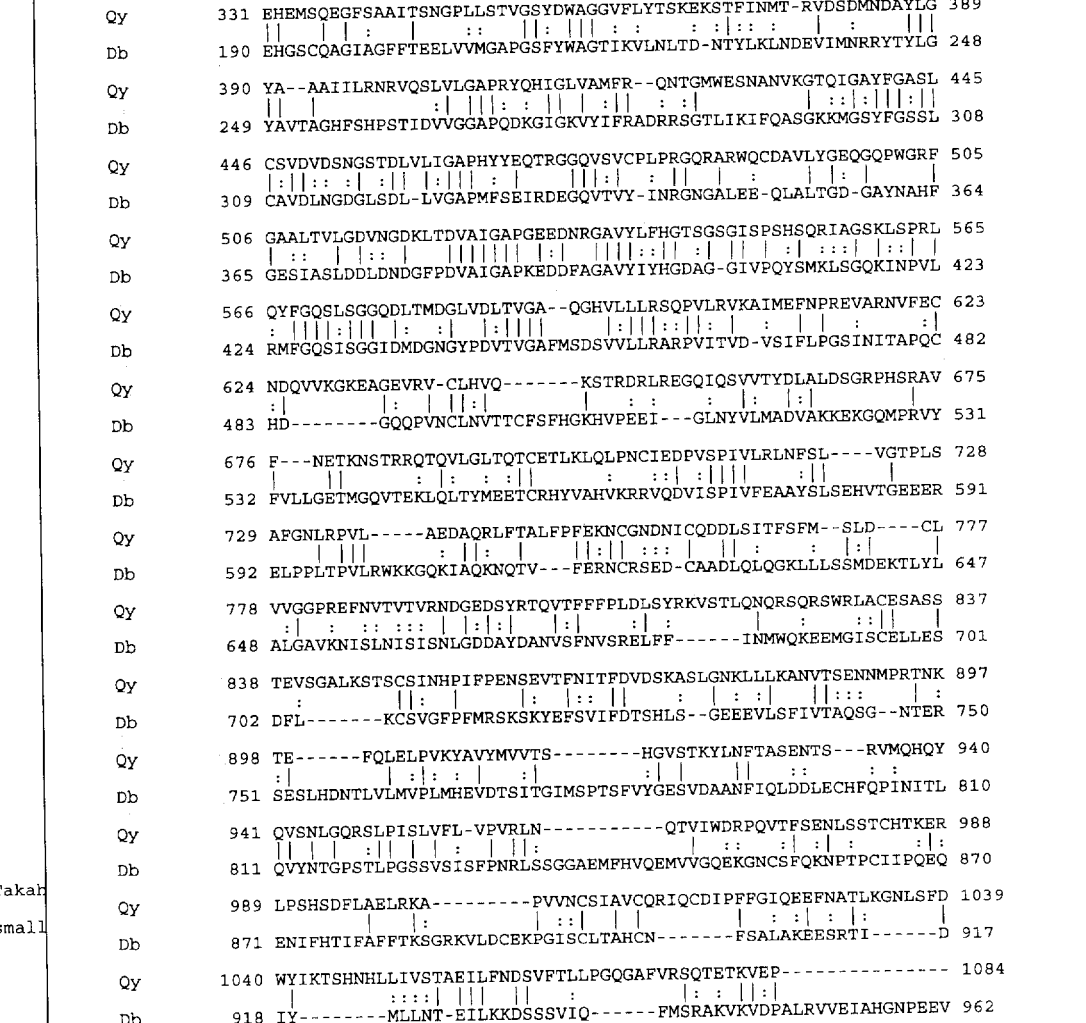
```
Qy          1 FNLDTENAMTFQ-ENARGFGQSVV-QLQGSR--VVVGAPQEIVAAN----QRGSLYQCDY 52

Db         40 YNVDTESALLYQGPHNTLFGYSVVLHSHGANRWLLVGAPTANWLANASVINPGAIYRCRI 99

Qy         53 STG---SCEPIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFL 109

Db        100 GKNPGQTCEQLQLGSP--------------NGEP----CG----KTCLEERDNQWL--- 133

Qy        110 FGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKT 169

Db        134 -GVTLSRQPGE--------------------NGSIVTCGHR--------------WKN 156

Qy        170 LFSLMQYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGA 229

Db        157 IFYI------------------------------------------------------- 160

Qy        230 RKNAFKILFLLTDGEKFGDPLGYEDVIPELDREGVIR-------YVLGFGDAFRSEKSRQ 282

Db        161 -KNENK----LPTGGCYGVP-------PDLRTELSKRIAPCYQDYVKKFGENFAS----- 203

Qy        283 ELNTVASKPPRDHVFQANNFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAA 342

Db        204 ----------------------------------------------------CQAGISSF 211

Qy        343 ITSNGPLLSTVGSYDWAGGVFLY---TSKEKSTFINMTRVDSDMNDAYLGYA--AAIILR 397

Db        212 YTKDLIVMGAPGSSYWTGSLFVYNITTNKYKAFLDKQNQVKF---GSYLGYSVGAGHFRS 268
```



RESULT 14
I58409
integrin alpha-9 chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I58409; A49459
R;Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah
Oncogene 9, 611-619, 1994
A;Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small
A;Reference number: I58409; MUID:94119603; PMID:8290272
A;Accession: I58409
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1035 <RES>
A;Cross-references: UNIPROT:Q13797; GB:D25303; NID:g464180; PIDN:BAA04984.1; PID:g533327

R;Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.
J. Cell Biol. 123, 1289-1297, 1993
A;Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partne
A;Reference number: A49459; MUID:94064789; PMID:8245132
A;Accession: A49459
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 30-1035 <PAL>
A;Cross-references: GB:L24158
C;Superfamily: integrin alpha-4 chain
C;Keywords: glycoprotein; metal binding; transmembrane protein
F;1-27/Domain: signal sequence #status predicted <SIG>

Query Match 10.7%; Score 630; DB 2; Length 1035;
Best Local Similarity 26.5%; Pred. No. 1.8e-35;
Matches 237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;



```
Qy       1085 ---FEVPN---PLPLIVG-----SSVGGLLLLALITAALYKLGFFKRQYKDMM 1126
                 ||  :   |  :||     | : |:|:  |:   |:|:|||:|:||:::
Db        963 TVVFEALHNLEPRGYVVGWIIAISLLVGILIFLLLAVLLWKMGFFRRRYKEII 1015


RESULT 15
T31437
integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C;Species: Lytechinus variegatus (variegated urchin)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31437
R;Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A;Description: Alpha SU2, a sea urchin integrin which binds laminin.
A;Reference number: Z21035
A;Accession: T31437
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1041 <HER>
A;Cross-references: UNIPROT:O76378; EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC2
A;Experimental source: developmental stage embryo
C;Function:
A;Description: binds laminin
C;Superfamily: integrin alpha-2b chain

  Query Match           10.5%;  Score 614.5;  DB 2;  Length 1041;
  Best Local Similarity  24.8%;  Pred. No. 2.2e-34;
  Matches  236;  Conservative  155;  Mismatches  314;  Indels  245;  Gaps   41;

Qy        336 QEGFSAAITSNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDS----DMNDAYLG 389
              | |||  | |:    ::   |||   | ::: :   :|  :  |:  :      :::| |
Db        181 QAGFSGIIFSDNSALVMGAPGSYYLQGQIYVQSLLNRSV-VQATQESNTGTYSFDNSYRG 239

Qy        390 YAAAI--ILRNRVQSLVLGAPRYQHI-GLVAMFRQNTGMWESNANVKGTQIGAYFGASLC 446
              |: |:     : ||  |:| || : : ||||:| ||   :     | |||| |||| |:
Db        240 YSLALGDFNGDGVQDYVVGTPRAESLMGLVAIFDQNLNQFN---QVMGTQIVAYFGYSVT 296

Qy        447 SVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPLPRGQRARWQCDAVLYGEQ-------- 498
              ||: :| : | :|:||| | :      |             ||:  ||     |
Db        297 VVDI-NNDTYDDLLVGAPMYMDDGPAIQ-----------RWEAGAVYVYLQNPDVGPGA 343

Qy        499 -----------GQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGS 546
                        ||   ||| ::  :|| | |   ||||||| | |: ||||::|| | :
Db        344 SNRLSLSSTLIGGQIRSRFGLSIASIGDSNQDGFNDVAIGAPYEGDDAGAVYIYHG-SAN 402

Qy        547 GISPSHSQRIAGSKLS-PRLQYFGQSLSGGQDLTMDGLVDLTVGAQ--GHVLLLRSQPVL 603
              |:  : :| :  | |    :  || || ||||:  :   || |||:     :|:|::||:
Db        403 GLKSTPAQVLTPSTLGHSGITTFGFSLQGGQDMDKNKYPDLLVGAESANTAVLIRTRPVV 462

Qy        604 RVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSV----- 658
              : | :   |                       : ::::  | : | :|: :
Db        463 SLDATLNTEP-----------------------IGINLENKTYE-LADGTMVTSFIAMT 497

Qy        659 -------------VTYDLALDSG-RPHSRAVFNETKNSTRRQTQVLGL-TQTCETLKLQ 702
                          ::| : :|||   : ||:| :   |   :|: | : || |: |:
Db        498 CFTYTGNYLPDHIDISYTVTVDSGIIANRRAMFVDNDMSEITKTRRLAVSTQFCDPLRAY 557

Qy        703 LPNCIEDPVSPIVLRLNFSLVGTPLSAFGN---------LRPVLAEDAQRLFTALFPFEK 753
              : | ||| ::|| : | : |          |          : |::         |     :
Db        558 VGNSIEDKLTPIKVTLQYDL--------NNDESRLQPHEILPIIDMATMSTQTKQVSIQN 609

Qy        754 NCGNDNICQDDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSY 813
              || | |||  || :| :  :|  :|:|  :|  : |:: | |||:::: :: ::|| | :
Db        610 NCVN-NICIPDLDVTVT-PNLPNIVIGQTQELTLDVSLNNRGEDAFQSSLSVYYPLGLQF 667

Qy        814 RKVSTLQNQRSQRSWRLACESASSTEVSGALKSTSCSINHPIFPEN------SEVTFNIT 867
              ::     :::   : : |   |        |:  :|   :|:  :|       :  || ::
Db        668 VRL----ERKANMDFSVTCSEDSD------LRIITCDTGNPMVGKNILEFGLTLSTFQVS 717

Qy        868 FDVDS-----KASLGNKLLLKANVTSENNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTK 922
              | ||      ||  |      |   | :  |         | :  | |  :  :
Db        718 GDKDSIEFYFKAESENS--EDPNTLENNELNMTVPVTVDCTLKLLSASYPEIVMYSTQED 775

Qy        923 YL-------NFTASENTSRVMQHQYQV-----SNLGQRSLPIS----------LVFLVPV 960
              |:         | : ::    || | |:|     || |: || |          | :|: :
Db        776 YVVPPFPAKNASEADIGMEVM-HLYEVRNTGSSNAGEVSLNIQWPQKNEDGEYLFYLLGI 834

Qy        961 RLNQTVIWDRPQ--------------VTFSENLSSTCHTKERLPSHSDFLAELRKAPVV 1005
                 : |      |                    | : :     | | |  :: ||:     |:
Db        835 MTEEGVTCQLTQGKANPEGVKLEPSTKAKLSNSTTQVSGRKRREPEVAEALAQTDN--VI 892

Qy       1006 NCSIAVCQRIQCDIPFFGIQEEFNAT------LKGNLSFDWYIKTSHNHLLIVSTAEILF 1059
                |:   |  | | |      :| ||:     : |    |                | |
Db        893 YCASDSCVLINCTI------DEINASKSKVVRILGRF---W--------------ERTF 928

Qy       1060 NDSVFTLLPGQGAFVRSQTETKVE--PFEVPNP-----------------LP------- 1092
                :|  | |   | : |     |:  |: :| |                 :|
Db        929 QKAVSELTPVVQATIASSASAAVKTIPYNIPLPRDFSDSTKASTLVTTEELVPPVTPIAW 988

Qy       1093 -LIVGSSVGGLLLLALITAALYKLGFFKR-------QYKDMMS--EGGPP 1132
                :|| | :||::|| :|   |:| |||:|       :|  : |  : |||
Db        989 WIIVVSVLGGIILLLIIILGLWKCGFFERKKPGEEKEYAPVASADKDGPP 1038


Search completed: November  9, 2004, 12:19:22
Job time : 34.25 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:46:42 ; Search time 136.75 Seconds
(without alignments)
4783.919 Million cell updates/sec

Title: US-09-902-481B-3
Perfect score: 5879
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5823.5	99.1	1152	1	ITAM_HUMAN	P11215 homo sapien
2	4447	75.6	1153	1	ITAM_MOUSE	P05555 mus musculu
3	4358	74.1	1151	2	Q9JI30	Q9ji30 rattus norv
4	3786.5	64.4	920	2	Q28984	Q28984 sus scrofa
5	3456	58.8	1163	1	ITAX_HUMAN	P20702 homo sapien
6	3388	57.6	1162	1	ITAD_HUMAN	Q13349 homo sapien
7	3306.5	56.2	1188	2	Q6KAS4	Q6kas4 mus musculu
8	3306.5	56.2	1188	2	BAD21383	Bad21383 mus muscu
9	3297.5	56.1	1169	1	ITAX_MOUSE	Q9qxh4 mus musculu
10	3225.5	54.9	1161	1	ITAD_RAT	Q9qye7 rattus norv
11	1561	26.6	1165	1	ITAL_BOVIN	P61625 bos taurus
12	1539	26.2	1166	2	Q6TYB8	Q6tyb8 bos taurus
13	1539	26.2	1166	2	AAQ90015	Aaq90015 bos tauru
14	1536.5	26.1	1170	1	ITAL_HUMAN	P20701 homo sapien
15	1516.5	25.8	1163	1	ITAL_MOUSE	P24063 mus musculu
16	1512.5	25.7	1161	2	Q9WTV4	Q9wtv4 mus musculu
17	1502	25.5	1160	2	Q9R200	Q9r200 mus musculu
18	1382	23.5	1196	2	Q98TF1	Q98tf1 cyprinus ca
19	1341.5	22.8	1086	2	Q96HB1	Q96hb1 homo sapien
20	1329.5	22.6	1187	2	Q98TF0	Q98tf0 cyprinus ca
21	1264	21.5	927	2	Q8HZV0	Q8hzv0 bos taurus
22	1142.5	19.4	1167	2	O88340	O88340 rattus norv
23	1136.5	19.3	1167	1	ITAE_MOUSE	Q60677 mus musculu
24	1128	19.2	1179	1	ITAE_HUMAN	P38570 homo sapien
25	1097	18.7	1167	2	O88341	O88341 rattus norv
26	1085.5	18.5	1151	1	ITA1_HUMAN	P56199 homo sapien
27	1079.5	18.4	1189	1	ITAH_HUMAN	Q9ukx5 homo sapien
28	1071	18.2	1170	1	ITA2_BOVIN	P53710 bos taurus
29	1057	18.0	1178	1	ITA2_MOUSE	Q62469 mus musculu
30	1057	18.0	1178	2	Q6P1C7	Q6p1c7 mus musculu
31	1057	18.0	1178	2	AAH65139	Aah65139 mus muscu
32	1054	17.9	1181	1	ITA2_HUMAN	P17301 homo sapien
33	1049	17.8	1180	1	ITA1_RAT	P18614 rattus norv
34	1042.5	17.7	1167	1	ITAG_HUMAN	O75578 homo sapien
35	1038.5	17.7	1188	1	ITAH_MOUSE	P61622 mus musculu
36	1038.5	17.7	1188	2	Q7TQC3	Q7tqc3 mus musculu
37	1034	17.6	1171	2	O42094	O42094 gallus gall
38	1020	17.3	1038	2	Q8BS01	Q8bs01 mus musculu
39	993	16.9	895	2	Q9WUF8	Q9wuf8 mus sp. itg
40	986.5	16.8	1160	2	Q8MKF4	Q8mkf4 felis silve
41	865	14.7	348	2	Q8TES5	Q8tes5 homo sapien
42	837	14.2	1332	2	Q9BPQ8	Q9bpq8 halocynthia
43	784	13.3	205	2	Q63001	Q63001 rattus norv
44	761.5	13.0	304	2	Q6PG66	Q6pg66 mus musculu
45	761.5	13.0	304	2	AAH57200	Aah57200 mus muscu

ALIGNMENTS

RESULT 1
ITAM_HUMAN
ID ITAM_HUMAN STANDARD; PRT; 1152 AA.
AC P11215;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha
DE subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)
DE (Neutrophil adherence receptor).
GN Name=ITGAM; Synonyms=CR3A, CD11B;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88315033; PubMed=2457584;
RA Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;
RT "The human leukocyte adhesion glycoprotein Mac-1 (complement receptor
RT type 3, CD11b) alpha subunit. Cloning, primary structure, and relation
RT to the integrins, von Willebrand factor and factor B.";
RL J. Biol. Chem. 263:12403-12411(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88190151; PubMed=2833753;
RA Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;
RT "Molecular cloning of the alpha subunit of human and guinea pig
RT leukocyte adhesion glycoprotein Mo1: chromosomal localization and
RT homology to the alpha subunits of integrins.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88257215; PubMed=2454931;
RA Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;
RT "Amino acid sequence of the alpha subunit of human leukocyte adhesion
RT receptor Mo1 (complement receptor type 3).";
RL J. Cell Biol. 106:2153-2158(1988).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93123748; PubMed=8419480;
RA Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;
RT "Structural analysis of the CD11b gene and phylogenetic analysis of
RT the alpha-integrin gene family demonstrate remarkable conservation of
RT genomic organization and suggest early diversification during
RT evolution.";
RL J. Immunol. 150:480-490(1993).
RN [5]
RP SEQUENCE OF 9-1153 FROM N.A.
RX MEDLINE=89098893; PubMed=2563162;
RA Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L.,
RA Roth G.J.;
RT "cDNA sequence for the alpha M subunit of the human neutrophil
RT adherence receptor indicates homology to integrin alpha subunits.";

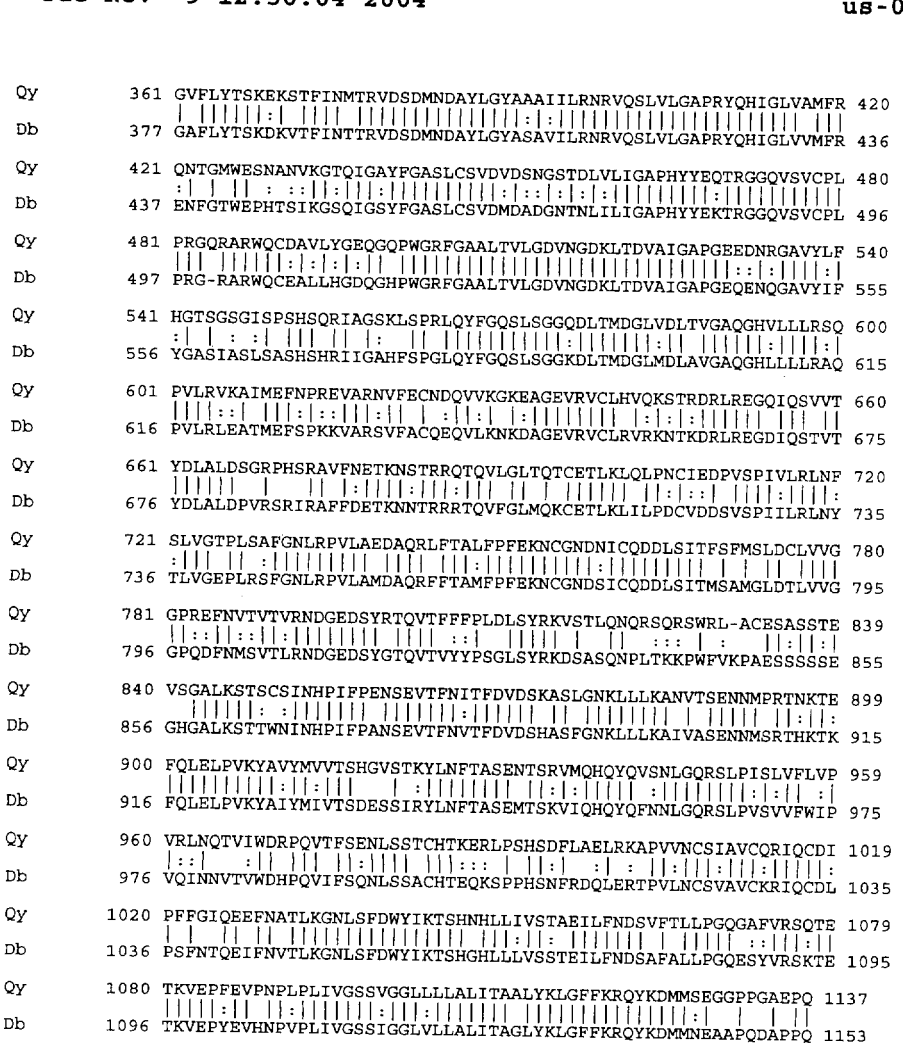
```
RL   Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).
RN   [6]
RP   SEQUENCE OF 1-9 FROM N.A.
RX   MEDLINE=92073318; PubMed=1683702;
RA   Shelley C.S., Arnaout M.A.;
RT   "The promoter of the CD11b gene directs myeloid-specific and
RT   developmentally regulated expression.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN   [7]
RP   SEQUENCE OF 1-9 FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=92144986; PubMed=1346576;
RA   Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT   "Characterization of the myeloid-specific CD11b promoter.";
RL   Blood 79:865-870(1992).
RN   [8]
RP   SEQUENCE OF 17-31.
RX   MEDLINE=87076671; PubMed=3539202;
RA   Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
RT   "N-terminal sequence of human leukocyte glycoprotein Mo1: conservation
RT   across species and homology to platelet IIb/IIIa.";
RL   Biochim. Biophys. Acta 874:368-371(1986).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX   MEDLINE=95171458; PubMed=7867070;
RA   Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT   "Crystal structure of the A domain from the alpha subunit of integrin
RT   CR3 (CD11b/CD18).";
RL   Cell 80:631-638(1995).
RN   [10]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX   MEDLINE=96363671; PubMed=8747460;
RA   Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
RT   "Two conformations of the integrin A-domain (I-domain): a pathway for
RT   activation?";
RL   Structure 3:1333-1340(1995).
RN   [11]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX   MEDLINE=98362595; PubMed=9687375;
RA   Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA   Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
RA   Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
RA   Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT   "Cation binding to the integrin CD11b I domain and activation model
RT   assessment.";
RL   Structure 6:923-935(1998).
RN   [12]
RP   3D-STRUCTURE MODELING OF 17-616.
RX   MEDLINE=98226734; PubMed=9560195;
RA   Oxvig C., Springer T.A.;
RT   "Experimental support for a beta-propeller domain in integrin alpha-
RT   subunits and a calcium binding site on its lower surface.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC   -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various
CC       adhesive interactions of monocytes, macrophages and granulocytes
CC       as well as in mediating the uptake of complement-coated particles.
CC       It is identical with CR-3, the receptor for the iC3b fragment of
CC       the third complement component. It probably recognizes the R-G-D
CC       peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for
CC       fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides
CC       of fibrinogen gamma chain.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-M
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC       granulocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; J03925; AAA59544.1; -.
DR   EMBL; M18044; AAA59491.1; -.
DR   EMBL; J04145; AAA59903.1; -.
DR   EMBL; S52227; AAB24821.1; -.
DR   EMBL; S52152; AAB24821.1; JOINED.
DR   EMBL; S52153; AAB24821.1; JOINED.
DR   EMBL; S52154; AAB24821.1; JOINED.
DR   EMBL; S52155; AAB24821.1; JOINED.
DR   EMBL; S52157; AAB24821.1; JOINED.
DR   EMBL; S52159; AAB24821.1; JOINED.
DR   EMBL; S52161; AAB24821.1; JOINED.
DR   EMBL; S52164; AAB24821.1; JOINED.
DR   EMBL; S52165; AAB24821.1; JOINED.
DR   EMBL; S52167; AAB24821.1; JOINED.
DR   EMBL; S52169; AAB24821.1; JOINED.
DR   EMBL; S52170; AAB24821.1; JOINED.
DR   EMBL; S52173; AAB24821.1; JOINED.
DR   EMBL; S52174; AAB24821.1; JOINED.
DR   EMBL; S52180; AAB24821.1; JOINED.
DR   EMBL; S52181; AAB24821.1; JOINED.
DR   EMBL; S52184; AAB24821.1; JOINED.
DR   EMBL; S52189; AAB24821.1; JOINED.
DR   EMBL; S52191; AAB24821.1; JOINED.
DR   EMBL; S52192; AAB24821.1; JOINED.
DR   EMBL; S52203; AAB24821.1; JOINED.
DR   EMBL; S52212; AAB24821.1; JOINED.
DR   EMBL; S52213; AAB24821.1; JOINED.
DR   EMBL; S52216; AAB24821.1; JOINED.
DR   EMBL; S52219; AAB24821.1; JOINED.
DR   EMBL; S52220; AAB24821.1; JOINED.
DR   EMBL; S52221; AAB24821.1; JOINED.
DR   EMBL; S52222; AAB24821.1; JOINED.
DR   EMBL; S52226; AAB24821.1; JOINED.
DR   EMBL; M76724; AAA58410.1; -.
DR   EMBL; M84477; AAA51960.1; -.
DR   PIR; A31108; RWHU1B.
DR   PDB; 1A8X; Model; @=17-1152.
DR   PDB; 1BHO; X-ray; 1/2=-.
DR   PDB; 1BHQ; X-ray; 1/2=-.
DR   PDB; 1IDN; X-ray; 1/2=-.
DR   PDB; 1IDO; X-ray; @=140-331.
DR   PDB; 1JLM; X-ray; @=143-334.
DR   PDB; 1M1U; X-ray; A=137-331.
DR   PDB; 1MF7; X-ray; A=144-337.
DR   PDB; 1N9Z; X-ray; A=140-335.
DR   PDB; 1NA5; X-ray; A=144-345.
DR   Genew; HGNC:6149; ITGAM.
DR   MIM; 120980; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0007155; P:cell adhesion; TAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
KW   Glycoprotein; Integrin; Magnesium; Receptor; Repeat; Signal;
KW   Transmembrane.
```

```
FT   SIGNAL       1    16
FT   CHAIN       17  1152       Integrin alpha-M.

  Query Match             99.1%;  Score 5823.5;  DB 1;  Length 1152;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 1123;  Conservative    8;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||| |||:||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||| |||||||||:| ||||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 2
ITAM_MOUSE
ID   ITAM_MOUSE     STANDARD;      PRT;  1153 AA.
AC   P05555; Q8CA73;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha
DE   subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).
GN   Name=Itgam;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=88312584; PubMed=3044779;
RA   Pytela R.;
RT   "Amino acid sequence of the murine Mac-1 alpha chain reveals homology
RT   with the integrin family and an additional domain related to von
RT   Willebrand factor.";
RL   EMBO J. 7:1371-1378(1988).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [3]
RP   SEQUENCE OF 11-45 FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Spleen;
RX   MEDLINE=86287312; PubMed=2942940;
RA   Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E.,
RA   Larson R.S., Roberts T.M., Springer T.A.;
RT   "A partial genomic DNA clone for the alpha subunit of the mouse
```

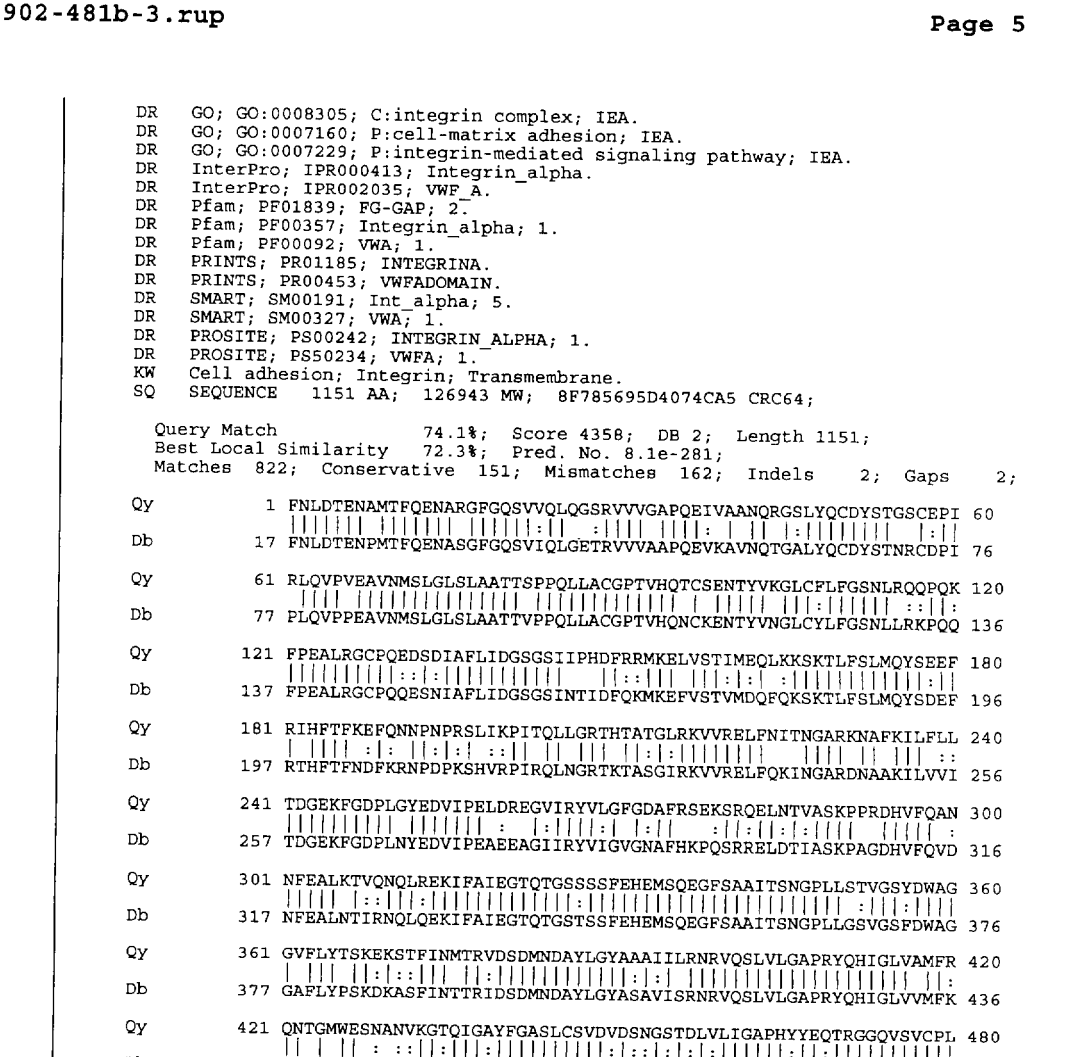
```
RT   complement receptor type 3 and cellular adhesion molecule Mac-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).
RN   [4]
RP   SEQUENCE OF 17-28.
RX   MEDLINE=85188276; PubMed=3887182;
RA   Springer T.A., Teplow D.B., Dreyer W.J.;
RT   "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
RT   glycoproteins and unexpected relation to leukocyte interferon.";
RL   Nature 314:540-542(1985).
CC   -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various
CC       adhesive interactions of monocytes, macrophages and granulocytes
CC       as well as in mediating the uptake of complement-coated particles.
CC       It is identical with CR-3, the receptor for the iC3b fragment of
CC       the third complement component. It probably recognizes the R-G-D
CC       peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for
CC       fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides
CC       of fibrinogen gamma chain. Alpha-M/beta-2 play a critical role in
CC       mast cell development and in immune complex-mediated
CC       glomerulonephritis. Mice expressing a null mutation of the alpha-M
CC       subunit gene demonstrate increase in neutrophil accumulation, in
CC       response to a impaired degranulation and phagocytosis, events that
CC       apparently accelerate apoptosis in neutrophils. These mice develop
CC       obesity.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta chain. Alpha-M
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P05555-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P05555-2; Sequence=VSP_010473;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC       granulocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X07640; CAA30479.1; -.
DR   EMBL; AK039444; BAC30350.1; -.
DR   EMBL; M14293; AAA39484.1; -.
DR   PIR; S00551; S00551.
DR   HSSP; P11215; 1BHQ.
DR   MGD; MGI:96607; Itgam.
DR   GO; GO:0009897; C:external side of plasma membrane; IDA.
DR   GO; GO:0007155; P:cell adhesion; IMP.
DR   GO; GO:0045123; P:cellular extravasation; IMP.
DR   GO; GO:0030593; P:neutrophil chemotaxis; IMP.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Alternative splicing; Calcium; Cell adhesion;
KW   Direct protein sequencing; Glycoprotein; Integrin; Receptor; Repeat;
KW   Signal; Transmembrane.
```

```
FT   SIGNAL       1     16
FT   CHAIN       17   1153       Integrin alpha-M.
FT   DOMAIN      17   1105       Extracellular (Potential).
FT   TRANSMEM  1106   1129       Potential.
FT   DOMAIN    1130   1153       Cytoplasmic (Potential).
FT   REPEAT      31     84       FG-GAP 1.
FT   REPEAT      85    163       FG-GAP 2.
FT   DOMAIN     164    350       VWFA.
FT   REPEAT     337    400       FG-GAP 3.
FT   REPEAT     401    452       FG-GAP 4.
FT   REPEAT     454    515       FG-GAP 5.
FT   REPEAT     517    575       FG-GAP 6.
FT   REPEAT     580    632       FG-GAP 7.
FT   CA_BIND    465    473       Potential.
FT   CA_BIND    529    537       Potential.
FT   CA_BIND    592    600       Potential.
FT   SITE      1132   1136       GFFKR motif.
FT   DISULFID    66     73       By similarity.
FT   DISULFID   105    123       By similarity.
FT   DISULFID   654    711       By similarity.
FT   DISULFID   770    776       By similarity.
FT   DISULFID   999   1023       By similarity.
FT   DISULFID  1028   1033       By similarity.
FT   CARBOHYD    58     58       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    86     86       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   391    391       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   696    696       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   734    734       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   772    772       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   801    801       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   881    881       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   907    907       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   941    941       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   980    980       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   994    994       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1022   1022       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1045   1045       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1051   1051       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1076   1076       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC   453    569       Missing (in isoform 2).
FT                               /FTId=VSP_010473.
FT   CONFLICT    37     37       N -> S (in Ref. 2).
FT   CONFLICT   683    683       V -> G (in Ref. 2).
SQ   SEQUENCE  1153 AA;  127480 MW;  178DB988AECB0343 CRC64;
```

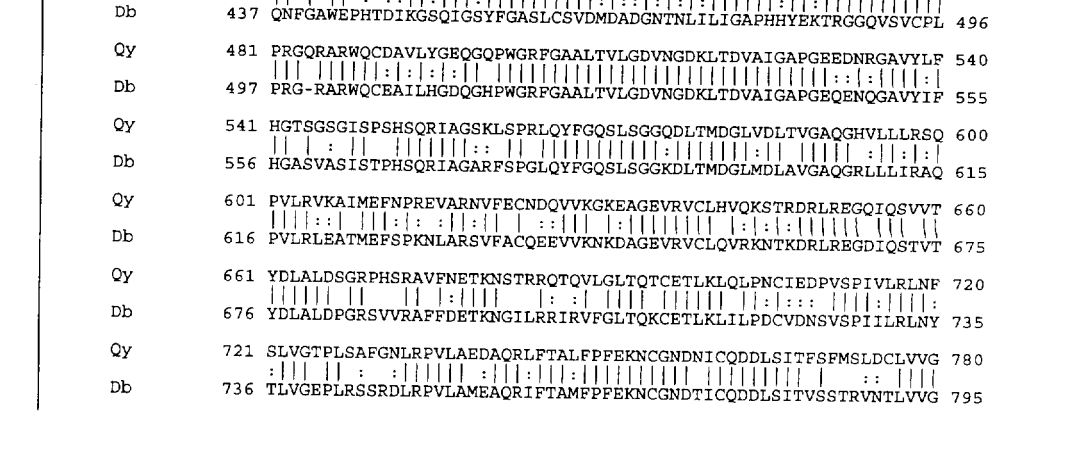
```
  Query Match           75.6%;  Score 4447;  DB 1;  Length 1153;
  Best Local Similarity 73.5%;  Pred. No. 9.7e-287;
  Matches  836;  Conservative  147;  Mismatches  153;  Indels    2;  Gaps    2;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||: ||||||:||||:|||| |: ||| |||| | || |:||||||| | ||
Db         17 FNLDTEHPMTFQENAKGFGQNVVQLGGTSVVVAAPQEAKAVNQTGALYQCDYSTSRCHPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||| :| | |||||||||||| | ||||| |||:|||||| : ||:
Db         77 PLQVPPEAVNMSLGLSLAVSTVPQQLLACGPTVHQNCKENTYVNGLCYLFGSNLLRPPQQ 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||| |||::||| ||||||||   ||::||| |||:||| ||||||||||||:||
Db        137 FPEALRECPQQESDIVFLIDGSGSINNIDFQKMKEFVSTVMEQFKKSKTLFSLMQYSDEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||| :|: ||:||| : || || ||| ||:|:|||||||: |||||:|| ||| ::
Db        197 RIHFTFNDFKRNPSPRSHVSPIKQLNGRTKTASGIRKVVRELFHKTNGARENAAKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||| |:||||| || |||||:| |:||   :||:||:|:|||| :|||| :
Db        257 TDGEKFGDPLDYKDVIPEADRAGVIRYVIGVGNAFNKPQSRRELDTIASKPAGEHVFQVD 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||| |:||||:|||||||||||:|||||||||||||||:|||||||| :|||:||||
Db        317 NFEALNTIQNQLQEKIFAIEGTQTGSTSSFEHEMSQEGFSASITSNGPLLGSVGSFDWAG 376
```



```
DR   GO; GO:0008305; C:integrin complex; IEA.
DR   GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR   GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 2.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Cell adhesion; Integrin; Transmembrane.
SQ   SEQUENCE   1151 AA;  126943 MW;  8F785695D4074CA5 CRC64;

  Query Match             74.1%;  Score 4358;  DB 2;  Length 1151;
  Best Local Similarity   72.3%;  Pred. No. 8.1e-281;
  Matches  822;  Conservative  151;  Mismatches  162;  Indels    2;  Gaps    2;
```



```
RESULT 3
Q9JI30
ID   Q9JI30      PRELIMINARY;      PRT;  1151 AA.
AC   Q9JI30;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Integrin beta 2 alpha subunit.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fathallah D.M. Sr., Zerria K. Jr.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR   EMBL; AF268593; AAF81280.1; -.
DR   HSSP; P11215; 1BHQ.
```



```
Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        796 DSRDFDVSVTLRNDGEDSYGTKVTCYYPSGLSYRKVSASQNQFSKKPWRVIAE-PSSSEG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 QGVLKSTIWDINHPIFPANSEVTFNVTFDVDSDASLGNKLLLKVVVTSENNVSRTDKTES 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        915 QLELPVKYAIYMVVTSGESSIKYLNFTASEMTSKVIQHQYQFNNLGQRSFPVSVVFWIPV 974

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        975 QINKVTIWDPPQVTFSQNLSSVCRTEQKSPSHSKFQDELERTPVLNCSVAVCKRIQCDIP 1034

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1035 SFNSKEIFNVTLQGNLLFDWYIETSHDHLLLVSTAEILFNDSAFALLPGQETFVKAQTET 1094

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1095 KVEPYTVHNPVPLIVGSSVGGLVLLALITAGLYKLGFFKRQYKDMMNEAGGQDGPPQ 1151
```

RESULT 4
Q28984
ID Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CD11b (Fragment).
GN Name=CD11b;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; U40072; AAB16869.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Cell adhesion; Integrin; Transmembrane.
FT NON_TER 1 1
FT NON_TER 920 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;

Query Match 64.4%; Score 3786.5; DB 2; Length 920;
Best Local Similarity 78.4%; Pred. No. 6.4e-243;
Matches 722; Conservative 85; Mismatches 113; Indels 1; Gaps 1;

```
Qy        118 PQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYS 177
Db          1 PQKFPEALRGCPQQESDIAFLIDGSGSINRLDFQRMKEFVSTVMGQFQKSKTLFALMQYS 60

Qy        178 EEFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKIL 237
Db         61 EDFYTHFTFNDFKRNPSPKLLVRPIRQLLGRTHTATGIRKVVRELFHSKSGARENALKIL 120

Qy        238 FLLTDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVF 297
Db        121 VVITDGEKFGDPLGYEDVIPEADRKGVIRYVIGVGDAFNSWKSREELNTIASKPCGDHVF 180

Qy        298 QANNFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYD 357
Db        181 QVNNFEAVKTIQNQLQEKTFAIEGTQTGSTSSFECEMSQEGFSAAITSNGPLLGAVGSFD 240

Qy        358 WAGGVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVA 417
Db        241 WAGGAFLHMPKDRVIFINTTRVDSDMNDAYLGYAVEVILRNQAQSLVLGAPRYQHTGLVV 300

Qy        418 MFRQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSV 477
Db        301 MFKQNSGAWEKNADIKGSQIGSYFGASLCSVDVNRDGSSDLVLIGAPHYYEQTRGGQVSV 360

Qy        478 CPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAV 537
Db        361 CPLPQG-RAKWQCRVILCGEQGHPWSRFGAALTALGDVNGDKLTDVAIGAPGEQDNRGAV 419

Qy        538 YLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLL 597
Db        420 YLFHGTSELGISPSHSQRIAGSQLSPRLQYFGQSLSGGQDLTMDGLMDLAVGAQGHVFLL 479

Qy        598 RSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQS 657
Db        480 RSQPVLRVEASMVFKPREVARNVYECRQQAAKTQIAGEVQVCLQVRKSTWDRLREGDTQS 539

Qy        658 VVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLR 717
Db        540 IITYDLALDPGRPHPRAVFEETKNNTRRQTQTLGLSRKCEHLALWLPDCVEDSVTPIVLR 599

Qy        718 LNFSLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCL 777
Db        600 LNFSLVGKPASSFGNLRPVLAVDAQRLFTALFPFEKNCGNDSICQDDLSITFSFMSLDTL 659

Qy        778 VVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASS 837
Db        660 VVGGPRDLKVTLTVRNQGEDSYRTQVTFFYPSDLSYRKVSTSQNQRSQRSWRLACESDVS 719

Qy        838 TEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNK 897
Db        720 TEESTALKSTSCSINHPIFPDNSEVTFNVTFDVNPDAFLGYKLLLKANVTSENNMPSSNK 779

Qy        898 TEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFL 957
Db        780 TEFQLELPVKYAVYVVVTSLEVSTKYFNFTASEKTRHVIEHQYQFNNLGQRKLPISVVFW 839

Qy        958 VPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQC 1017
Db        840 VPVRLNRVTVWDQPQVTFSQNLSRSCSTEEIGPRHSDFLEKLQKTPVLNCSIAVCQKIQC 899

Qy       1018 DIPFFGIQEEFNATLKGNLSF 1038
Db        900 DIPSFGIQEELKVTLKGNLSF 920
```

RESULT 5
ITAX_HUMAN
ID ITAX_HUMAN STANDARD; PRT; 1163 AA.
AC P20702; Q8IVA6;
DT 01-FEB-1991 (Rel. 17, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
GN Name=ITGAX; Synonyms=CD11C;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```
RX   MEDLINE=88166645; PubMed=3327687;
RA   Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT   "cDNA cloning and complete primary structure of the alpha subunit of a
RT   leukocyte adhesion glycoprotein, p150,95.";
RL   EMBO J. 6:4023-4028(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90153906; PubMed=2303426;
RA   Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT   "Genomic structure of an integrin alpha subunit, the leukocyte p150,95
RT   molecule.";
RL   J. Biol. Chem. 265:2782-2788(1990).
RN   [3]
RP   ERRATUM.
RA   Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL   J. Biol. Chem. 265:12750-12751(1990).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 20-43.
RX   MEDLINE=87167596; PubMed=3549901;
RA   Miller L.J., Wiebe M., Springer T.A.;
RT   "Purification and alpha subunit N-terminal sequences of human Mac-1
RT   and p150,95 leukocyte adhesion proteins.";
RL   J. Immunol. 138:2381-2383(1987).
CC   -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
CC       recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
CC       interaction during inflammatory responses. It is especially
CC       important in monocyte adhesion and chemotaxis.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC       granulocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M81695; AAA59180.1; -.
DR   EMBL; M29165; -; NOT_ANNOTATED_CDS.
DR   EMBL; M29487; AAA51620.1; ALT_SEQ.
DR   EMBL; M29482; AAA51620.1; JOINED.
DR   EMBL; M29483; AAA51620.1; JOINED.
DR   EMBL; M29484; AAA51620.1; JOINED.
DR   EMBL; M29485; AAA51620.1; JOINED.
DR   EMBL; M29486; AAA51620.1; JOINED.
DR   EMBL; BC038237; AAH38237.1; -.
DR   PIR; A36584; RWHU1C.
DR   PDB; 1N3Y; X-ray; A=141-338.
DR   Genew; HGNC:6152; ITGAX.
DR   MIM; 151510; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0004872; F:receptor activity; TAS.
DR   GO; GO:0007155; P:cell adhesion; TAS.
DR   GO; GO:0009887; P:organogenesis; TAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
KW   Glycoprotein; Integrin; Magnesium; Polymorphism; Receptor; Repeat;
KW   Signal; Transmembrane.
FT   SIGNAL       1     19
FT   CHAIN       20   1163       Integrin alpha-X.
FT   DOMAIN      20   1107       Extracellular (Potential).
FT   TRANSMEM  1108   1128       Potential.
FT   DOMAIN    1129   1163       Cytoplasmic (Potential).
FT   REPEAT      34     87       FG-GAP 1.
FT   REPEAT      88      ?       FG-GAP 2.
FT   DOMAIN     165    351       VWFA.
FT   REPEAT       ?    401       FG-GAP 3.
FT   REPEAT     402    453       FG-GAP 4.
FT   REPEAT     455    517       FG-GAP 5.
FT   REPEAT     518    576       FG-GAP 6.
FT   REPEAT     581    633       FG-GAP 7.
FT   CA_BIND    466    474       Potential.
FT   CA_BIND    530    538       Potential.
FT   CA_BIND    593    601       Potential.
FT   SITE      1131   1135       GFFKR motif.
FT   DISULFID    69     76       By similarity.
FT   DISULFID   108    126       By similarity.
FT   DISULFID   655    712       By similarity.
FT   DISULFID   771    777       By similarity.
FT   DISULFID   848    863       By similarity.
FT   DISULFID   998   1022       By similarity.
FT   DISULFID  1027   1032       By similarity.
FT   CARBOHYD    61     61       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   392    392       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   697    697       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   735    735       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   899    899       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   939    939       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1050   1050       N-linked (GlcNAc. . .) (Potential).
FT   VARIANT     48     48       W -> R (in dbSNP:11574633).
FT                               /FTId=VAR_018672.
FT   CONFLICT   209    209       T -> S (in Ref. 4).
FT   CONFLICT   251    251       T -> A (in Ref. 4).
FT   CONFLICT   469    469       T -> S (in Ref. 4).
FT   CONFLICT   490    490       G -> A (in Ref. 2).
FT   CONFLICT   547    547       E -> K (in Ref. 4).
FT   CONFLICT   756    756       D -> L (in Ref. 1).
FT   CONFLICT   819    819       I -> V (in Ref. 4).
FT   CONFLICT  1161   1163       SEK -> TPHYPQDNV (in Ref. 4).
FT   STRAND     150    157
```

```
FT   TURN        160    161
FT   HELIX       164    178
FT   TURN        179    180
FT   TURN        183    185
FT   STRAND      186    193
FT   STRAND      197    201
FT   HELIX       203    208
FT   HELIX       212    216
FT   TURN        217    218
FT   STRAND      226    226
FT   HELIX       228    236
FT   TURN        237    240
FT   HELIX       242    244
FT   TURN        245    245
FT   TURN        248    249
FT   STRAND      251    258
FT   STRAND      263    263
FT   HELIX       269    278
FT   TURN        279    280
FT   STRAND      282    288
FT   HELIX       290    293
FT   TURN        296    297
FT   HELIX       298    304
FT   HELIX       310    312
FT   STRAND      313    316
FT   HELIX       319    325
FT   HELIX       326    334
FT   TURN        335    335
SQ   SEQUENCE   1163 AA;  127886 MW;  83658A13B5C5DE8F CRC64;

  Query Match             58.8%;  Score 3456;  DB 1;  Length 1163;
  Best Local Similarity   60.6%;  Pred. No. 9.3e-221;
  Matches  684;  Conservative  143;  Mismatches  296;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||| |: :: ||| |||| | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||||:|||| |||||||||| | | |: ||||| |    | |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || |||||||||    :| |   |  :: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              : ||||:||:   || ||:  : || | |:||| :: ||  ||: : |||::| ||| ::
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||:| || | |:|||| | |:||| :| | ||::  | :||| :|||| ::|:|:
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| :|||:|||||||||:| |||||| ||:|||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |   ||||||||||||| |    :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:||||||||||||::|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  |||||||||||:|| |||||||:|||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :||| :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :||||   ||    | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || |||||:|| |||| |||  ||||||| |:||||:| |:|||  |   |:||
Db        737 TLVGKPLLAFRNLRPMLAADAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |  |||| ||| ||   :::||  |||| || ||::||| |||:|||| |||:||  |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||||| ||:||   ||||||: | |  | |  |:|||:||||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| |  | :::  |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :      : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143


RESULT 6
ITAD_HUMAN
ID   ITAD_HUMAN     STANDARD;      PRT;  1162 AA.
AC   Q13349; Q15575; Q15576;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN   Name=ITGAD;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=96111956; PubMed=8777714;
RA   Van der Vieren M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA   Staunton D.E., Gallatin W.M.;
RT   "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
RT   3.";
RL   Immunity 3:683-690(1995).
RN   [2]
RP   SEQUENCE OF 1-235 FROM N.A.
RX   MEDLINE=20187620; PubMed=10722744;
RA   Noti J.D., Johnson A.K., Dillon J.D.;
RT   "Structural and functional characterization of the leukocyte integrin
RT   gene CD11d. Essential role of Sp1 and Sp3.";
RL   J. Biol. Chem. 275:8959-8969(2000).
RN   [3]
RP   SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX   MEDLINE=96257236; PubMed=8666289;
RA   Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
RT   "Cloning and chromosomal localization of a novel gene-encoding a human
RT   beta 2-integrin alpha subunit.";
```

```
RL   Gene 171:291-294(1996).
RN   [4]
RP   INTERACTION WITH VCAM1.
RX   MEDLINE=99059842; PubMed=9841932;
RA   Grayson M.H., Van der Vieren M., Sterbinsky S.A., Michael Gallatin W.,
RA   Hoffman P.A., Staunton D.E., Bochner B.S.;
RT   "alphadbeta2 integrin is expressed on human eosinophils and functions
RT   as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-
RT   1).";
RL   J. Exp. Med. 188:2187-2191(1998).
RN   [5]
RP   INTERACTION WITH VCAM1.
RX   MEDLINE=99370002; PubMed=10438935;
RA   Van der Vieren M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
RA   Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RT   "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
RT   binding interface between I domain and VCAM-1.";
RL   J. Immunol. 163:1984-1990(1999).
CC   -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC       VCAM1. May play a role in the atherosclerotic process such as
CC       clearing lipoproteins from plaques and in phagocytosis of blood-
CC       borne pathogens, particulate matter, and senescent erythrocytes
CC       from the blood.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell
CC       lines and subsets of peripheral blood leukocytes and strongly on
CC       tissue-specialized cells, including macrophages foam cells within
CC       atherosclerotic plaques, and on splenic red pulp macrophages.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U37028; AAB38547.1; -.
DR   EMBL; U40274; AAB60634.1; -.
DR   EMBL; U40275; AAB60635.1; -.
DR   EMBL; U40276; AAB60636.1; -.
DR   EMBL; U40277; AAB60637.1; -.
DR   EMBL; U40279; AAB60638.1; -.
DR   EMBL; U40278; AAB60638.1; JOINED.
DR   EMBL; AF187881; AAF62875.1; -.
DR   HSSP; P11215; 1BHQ.
DR   Genew; HGNC:6146; ITGAD.
DR   MIM; 602453; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0016337; P:cell-cell adhesion; NAS.
DR   GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     17       Potential.
FT   CHAIN        18   1162       Integrin alpha-D.
FT   DOMAIN       18   1100       Extracellular (Potential).
FT   TRANSMEM   1101   1121       Potential.
FT   DOMAIN     1122   1162       Cytoplasmic (Potential).
FT   REPEAT       32     85       FG-GAP 1.
FT   REPEAT       86      ?       FG-GAP 2.
FT   DOMAIN      150    332       VWFA.
FT   REPEAT      350    400       FG-GAP 3.
FT   REPEAT      401    452       FG-GAP 4.
FT   REPEAT      454    516       FG-GAP 5.
FT   REPEAT      518    576       FG-GAP 6.
FT   REPEAT      581    633       FG-GAP 7.
FT   CA_BIND     465    473       Potential.
FT   CA_BIND     530    538       Potential.
FT   CA_BIND     593    601       Potential.
FT   SITE       1127   1131       GFFKR motif.
FT   DISULFID     67     74       By similarity.
FT   DISULFID    106    124       By similarity.
FT   DISULFID    655    710       By similarity.
FT   DISULFID    769    775       By similarity.
FT   DISULFID    846    861       By similarity.
FT   DISULFID    994   1018       By similarity.
FT   DISULFID   1023   1028       By similarity.
FT   CARBOHYD     59     59       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     87     87       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     99     99       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    391    391       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    691    691       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    733    733       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    873    873       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    957    957       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1046   1046       N-linked (GlcNAc. . .) (Potential).
FT   CONFLICT    500    500       Missing (in Ref. 2).
FT   CONFLICT    515    518       GHPW -> ATP (in Ref. 2).
FT   CONFLICT    825    825       L -> V (in Ref. 2).
FT   CONFLICT    984    984       V -> A (in Ref. 2).
SQ   SEQUENCE   1162 AA;  126885 MW;  F296A1A35455D77D CRC64;

  Query Match             57.6%;  Score 3388;  DB 1;  Length 1162;
  Best Local Similarity   58.9%;  Pred. No. 3.1e-216;
  Matches  665;  Conservative  168;  Mismatches  288;  Indels    8;  Gaps    6;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    |||:| |||||||  |||:|||||  |:|||||  | || |   :|| |:||
Db         18 FNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQTGRLYDCAAATGMCQPI 77

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              | :  |||||||||:|||:|:  :|||||||:|: | ||:| || | | ||   :  |
Db         78 PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-EIIQT 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
                |:|   || :: || |||||||||  :|| :||  |  :| | : : |||:|||||
Db        137 VPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFEGTDTLFALMQYSNLL 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              :|||||  :|: :|: :||: || || | | ||||:  || :||:  |||||:| ||| ::
Db        197 KIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHHKNGARKSAKKILIVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||:|: ||| | |||| :: |:||| :| | ||:   :||||||::| ||:||||: :
Db        257 TDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELNTISSAPPQDHVFKVD 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              || || ::| ||:|||:|:|||| :||||:||||||||| |:| :|  | |||: |:|
Db        317 NFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGSFSWSG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |   ||:|||||||||| |   :|
Db        377 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVLGAPRYQHTGKAVIFT 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:|||||||||||||:|||||:|||||||||||||||||||||
Db        437 QVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPHYYEQTRGGQVSVCPL 496
```

```
Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        497 PRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAIGAPGEQENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLLRSL 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 674

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        675 FDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCVEDVVSPIILHLNF 734

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        735 SLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDLGVTLSFSGLQTLTVG 794

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        795 SSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 853

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        854 EG-LRSSRCSVNHPIFHEGSNGTFIVTFDVSYKATLGDRMLMRASASSENNKASSSKATF 912

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNF-TASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        913 QLELPVKYAVYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQRDLAISINFWVP 972

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        973 VLLNGVAVWDVVMEAPSQSL--PCVSERKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1030

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1031 PSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQLPGQEAFMRAQME 1090

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1091 MVLEEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRHYKEMLED 1139
```

```
RESULT 7
Q6KAS4
ID   Q6KAS4      PRELIMINARY;      PRT;  1188 AA.
AC   Q6KAS4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MFLJ00114 protein (Fragment).
GN   Name=mFLJ00114;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA   Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT   "Prediction of the Coding Sequences of Mouse Homologues of FLJ Genes:
RT   The Complete Nucleotide Sequences of 110 Mouse FLJ-Homologous cDNAs
RT   Identified by Screening of Terminal Sequences of cDNA Clones Randomly
RT   Sampled from Size-Fractionated Libraries.";
RL   DNA Res. 11:167-180(2004).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR   EMBL; AK131133; BAD21383.1; -.
DR   GO; GO:0009897; C:external side of plasma membrane; IDA.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; UNKNOWN_1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Cell adhesion; Integrin; Transmembrane.
FT   NON_TER      1       1
SQ   SEQUENCE   1188 AA;  131248 MW;  B8D93107BDBB4178 CRC64;
```

```
  Query Match             56.2%;  Score 3306.5;  DB 2;  Length 1188;
  Best Local Similarity   56.4%;  Pred. No. 8.6e-211;
  Matches  643;  Conservative  172;  Mismatches  306;  Indels   19;  Gaps    7;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         39 FNLDAEKPTHFHMDGAEFGHSVLQYDSSWVVVGAPKEIKATNQIGGLYKCGYHTGNCEPI 98

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         99 SLQVPPEAVNMSLGLSLAAATNPSWLLACGPTVHHTCRENIYLTGLCFLLSSSFKQS-QN 157

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
Db        158 FPTAQQECPKQDQDIVFLIDGSGSISSTDFEKMLDFVKAVMSQLQRPSTRFSLMQFSDYF 217

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db        218 RVHFTFNNFISTSSPLSLLDSVRQLRGYTYTASAIKHVITELFTTQSGARQDATKVLIVI 277

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db        278 TDGRKQGDNLSYDSVIPMAEAASIIRYAIGVGKAFYNEHSKQELKAIASMPSHEYVFSVE 337

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        338 NFDALKDIENQLKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTPDGPVLGAVGSFSWSG 397

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        398 GAFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVVIFT 457

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        458 QESRHWRPKSEVRGTQIGSYFGASLCSVDMDRDGSTDLVLIGVPHYYEHTRGGQVSVCPM 517

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        518 P-GVGRRWHCGTTLHGEQGHPWGRFGAALTVLGDVNGDSLADVAIGAPGEEENRGAVYIF 576

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        577 HGASRQDIAPSPSQRISASQIPSRIQYFGQSLSGGQDLTRDGLVDLAVGSKGRVLLLRTR 636

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        637 PILRVSPTVHFTPAEISRSVFECQEQVAPEQTLSDATVCLHIHESPKTQL--GDLRSTVT 694

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        695 FDLALDHGRLSTRAIFKETKTRALTRVKTLGLNKHCESVKLLLPACVEDSVTPITLRLNF 754

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        755 SLVGVPISSLQNLQPMLAVDDQTYFTASLPFEKNCGADHICQDDLSVVFGFPDLKTLVVG 814

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQ----------NQRSQRSWR 829
Db        815 SDLELNVDVTVSNDGEDSYGTTVTLFYPVGLSFRRVAEGQVFLRKKEDQQWQRRGQHSLH 874
```

```
Qy        830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSE 889
Db        875 LMCD--STPDRSQGLWSTSCSSRHVIFRGGSQMTFLVTFDVSPKAELGDRLLLRARVGSE 932

Qy        890 NNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASE-NTSRVMQHQYQVSNLGQR 948
Db        933 NNVPGTPKTTFQLELPVKYAVYTMISSHDQFTKYLNFSTSEKEKTSVVEHRFQVNNLGQR 992

Qy        949 SLPISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCS 1008
Db        993 DVPVSINFWVPIELKGEAVW-TVMVSHPQNPLTQCYRNRLKPTQFDFLTHMQKSPVLDCS 1051

Qy       1009 IAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLP 1068
Db       1052 IADCLHLRCDIPSLGILDELYFILKGNLSFGWISQTLQKKVLLLSEAEITFNTSVYSQLP 1111

Qy       1069 GQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1112 GQEAFLRAQTKTVLEMYKVHNPVPLIVGSSVGGLLLLAIITAILYKAGFFKRQYKEMLEE 1171
```

RESULT 8
BAD21383
ID BAD21383 PRELIMINARY; PRT; 1188 AA.
AC BAD21383;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE MFLJ00114 protein (Fragment).
GN MFLJ00114.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Gene:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-homologous cDNAs
RT Identified by Screening of Terminal sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK131133; BAD21383.1; -.
FT NON_TER 1 1
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDBB4178 CRC64;

Query Match 56.2%; Score 3306.5; DB 2; Length 1188;
Best Local Similarity 56.4%; Pred. No. 8.6e-211;
Matches 643; Conservative 172; Mismatches 306; Indels 19; Gaps 7;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         39 FNLDAEKPTHFHMDGAEFGHSVLQYDSSWVVVGAPKEIKATNQIGGLYKCGYHTGNCEPI 98

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         99 SLQVPPEAVNMSLGLSLAAATNPSWLLACGPTVHHTCRENIYLTGLCFLLSSSFKQS-QN 157

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
Db        158 FPTAQQECPKQDQDIVFLIDGSGSISSTDFEKMLDFVKAVMSQLQRPSTRFSLMQFSDYF 217

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db        218 RVHFTFNNFISTSSPLSLLDSVRQLRGYTYTASAIKHVITELFTTQSGARQDATKVLIVI 277

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db        278 TDGRKQGDNLSYDSVIPMAEAASIIRYAIGVGKAFYNEHSKQELKAIASMPSHEYVFSVE 337

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        338 NFDALKDIENQLKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTPDGPVLGAVGSFSWSG 397

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        398 GAFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVVIFT 457

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        458 QESRHWRPKSEVRGTQIGSYFGASLCSVDMDRDGSTDLVLIGVPHYYEHTRGGQVSVCPM 517

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        518 P-GVGRRWHCGTTLHGEQGHPWGRFGAALTVLGDVNGDSLADVAIGAPGEEENRGAVYIF 576

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        577 HGASRQDIAPSPSQRISASQIPSRIQYFGQSLSGGQDLTRDGLVDLAVGSKGRVLLLRTR 636

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        637 PILRVSPTVHFTPAEISRSVFECQEQVAPEQTLSDATVCLHIHESPKTQL--GDLRSTVT 694

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        695 FDLALDHGRLSTRAIFKETKTRALTRVKTLGLNKHCESVKLLLPACVEDSVTPITLRLNF 754

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        755 SLVGVPISSLQNLQPMLAVDDQTYFTASLPFEKNCGADHICQDDLSVVFGFPDLKTLVVG 814

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQ----------NQRSQRSWR 829
Db        815 SDLELNVDVTVSNDGEDSYGTTVTLFYPVGLSFRRVAEGQVFLRKKEDQQWQRRGQHSLH 874

Qy        830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSE 889
Db        875 LMCD--STPDRSQGLWSTSCSSRHVIFRGGSQMTFLVTFDVSPKAELGDRLLLRARVGSE 932

Qy        890 NNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASE-NTSRVMQHQYQVSNLGQR 948
Db        933 NNVPGTPKTTFQLELPVKYAVYTMISSHDQFTKYLNFSTSEKEKTSVVEHRFQVNNLGQR 992

Qy        949 SLPISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCS 1008
Db        993 DVPVSINFWVPIELKGEAVW-TVMVSHPQNPLTQCYRNRLKPTQFDFLTHMQKSPVLDCS 1051

Qy       1009 IAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLP 1068
Db       1052 IADCLHLRCDIPSLGILDELYFILKGNLSFGWISQTLQKKVLLLSEAEITFNTSVYSQLP 1111

Qy       1069 GQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1112 GQEAFLRAQTKTVLEMYKVHNPVPLIVGSSVGGLLLLAIITAILYKAGFFKRQYKEMLEE 1171
```

RESULT 9
ITAX_MOUSE
ID ITAX_MOUSE STANDARD; PRT; 1169 AA.
AC Q9QXH4;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c).
GN Name=Itgax;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Dendritic cell;

```
RA   Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RA   Tsuchiya H.;
RT   "Isolation of genes selectively expressed by dendritic cells.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
CC       recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
CC       interaction during inflammatory responses. It is especially
CC       important in monocyte adhesion and chemotaxis (By similarity).
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
CC       associates with beta-2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF211864; AAF23492.1; -.
DR   HSSP; P20702; 1N3Y.
DR   MGD; MGI:96609; Itgax.
DR   GO; GO:0009897; C:external side of plasma membrane; IDA.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     19       Potential.
FT   CHAIN        20   1169       Integrin alpha-X.
FT   DOMAIN       20   1116       Extracellular (Potential).
FT   TRANSMEM   1117   1137       Potential.
FT   DOMAIN     1138   1169       Cytoplasmic (Potential).
FT   REPEAT       34     87       FG-GAP 1.
FT   REPEAT       88      ?       FG-GAP 2.
FT   DOMAIN      152    330       VWFA.
FT   REPEAT        ?    402       FG-GAP 3.
FT   REPEAT      403    454       FG-GAP 4.
FT   REPEAT      456    518       FG-GAP 5.
FT   REPEAT      519    577       FG-GAP 6.
FT   REPEAT      582    634       FG-GAP 7.
FT   CA_BIND     467    475       Potential.
FT   CA_BIND     531    539       Potential.
FT   CA_BIND     594    602       Potential.
FT   SITE       1140   1144       GFFKR motif.
FT   DISULFID     69     76       By similarity.
FT   DISULFID    108    126       By similarity.
FT   DISULFID    656    711       By similarity.
FT   DISULFID    770    776       By similarity.
FT   DISULFID    858    873       By similarity.
FT   DISULFID   1007   1031       By similarity.
FT   DISULFID   1036   1041       By similarity.
FT   CARBOHYD     89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    267    267       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    393    393       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    734    734       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    949    949       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1059   1059       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1084   1084       N-linked (GlcNAc. . .) (Potential).
```

```
SQ   SEQUENCE   1169 AA;  129150 MW;  C616412033C219A6 CRC64;

  Query Match             56.1%;  Score 3297.5;  DB 1;  Length 1169;
  Best Local Similarity   56.2%;  Pred. No. 3.3e-210;
  Matches  641;  Conservative  174;  Mismatches  306;  Indels   19;  Gaps    7;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    | :   || ||:|   | |||||:|| | || | ||:| | ||:|||| 
Db         20 FNLDAEKLTHFHMDGAEFGHSVLQYDSSWVVVGAPKEIKATNQIGGLYKCGYHTGNCEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||:|||||||| |:|  |||||||||| || || |: |||||  |: :| | 
Db         80 SLQVPPEAVNISLGLSLAAATNPSWLLACGPTVHHTCRENIYLTGLCFLLSSSFKQS-QN 138

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              || | : ||::| || |||||||||   || :| : |  :| ||::  | |||||:|: | 
Db        139 FPTAQQECPKQDQDIVFLIDGSGSISSTDFEKMLDFVKAVMSQLQRPSTRFSLMQFSDYF 198

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |:|||  | :  :| ||:  : || | |:||: :: |: |||   :|||::| |:| :: 
Db        199 RVHFTFNNFISTSSPLSLLGSVRQLRGYTYTASAIKHVITELFTTQSGARQDATKVLIVI 258

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              ||| | || | |: |||  :  :|||:| | || :| |:|||  :|| |  ::||
Db        259 TDGRKQGDNLSYDSVIPMAEAASIIRYAIGVGKAFYNEHSKQELKAIASMPSHEYVFSVE 318

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||:||| ::|||:|||||||||:| |||:|| |||||||||  | :||:|  |||: |:| 
Db        319 NFDALKDIENQLKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTPDGPVLGAVGSFSWSG 378

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | ||| |  : |||||:: : || |||||: |:      | ||:|||||:|| | | :| 
Db        379 GAFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVVIFT 438

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   : |:|||||:|||||||||||:| :|||||||| ||||| ||||||||||: 
Db        439 QESRHWRPKSEVRGTQIGSYFGASLCSVDMDRDGSTDLVLIGVPHYYEHTRGGQVSVCPM 498

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              | |  :|| |   |:|||| |||||||||||||||||| | ||||||||||:|||||:| 
Db        499 P-GVGSRWHCGTTLHGEQGHPWGRFGAALTVLGDVNGDSLADVAIGAPGEEENRGAVYIF 557

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || |   |:|| |||: |::  |:|||||||||||||| |||||| ||::| |||||::
Db        558 HGASRQDIAPSPSQRISASQIPSRIQYFGQSLSGGQDLTRDGLVDLAVGSKGRVLLLRTR 617

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |:|||   : | | |::|:|||| :||  :  :  ||||: :| : :|  | ::| ||
Db        618 PILRVSPTVHFTPAEISRSVFECQEQVAPEQTLSDATVCLHIHESPKTQL--GDLRSTVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              :||||| ||  :||:| |||     : : ||| : ||:|| || |:|| |:|| ||||| 
Db        676 FDLALDHGRLSTRAIFKETKTRALTRVKTLGLNKHCESVKLLLPACVEDSVTPITLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||| |:|:  ||:|:|| | |  ||| |||||| |:||||||: | |  |  ||||
Db        736 SLVGVPISSLQNLQPMLAVDDQTYFTASLPFEKNCGADHICQDDLSVVFGFPDLKTLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQ----------NQRSQRSWR 829
              | || ||| |||||| | || |:|: ||:|:|:  |           :| | |
Db        796 SDLELNVDVTVSNDGEDSYGTTVTLFYPVGLSFRRVAEGQVFLRKKEDQQWQRRGQHSLH 855

Qy        830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSE 889
              | |:  |: : |  | ||||  | ||   |::|| :|||  || ||::|||:| | ||
Db        856 LMCD--STPDRSQGLWSTSCSSRHVIFRGGSQMTFLVTFDVSPKAELGDRLLLRARVGSE 913

Qy        890 NNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASE-NTSRVMQHQYQVSNLGQR 948
              ||:| | || ||||||||||||| :::||   ||||||: ||   : |::|::||:|||||
Db        914 NNVPGTPKTTFQLELPVKYAVYTMISSHDQFTKYLNFSTSEKEKTSVVEHRFQVNNLGQR 973

Qy        949 SLPISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCS 1008
              :|:|: | ||: |    :|    |:  :|  : |:     |: | |  ::|:||::||
Db        974 DVPVSINFWVPIELKGEAVW-TVMVSHPQNPLTQCYRNRLKPTQFDLLTHMQKSPVLDCS 1032
```

```
Qy       1009 IAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLP 1068
Db       1033 IADCLHLRCDIPSLGILDELYFILKGNLSFGWISQTLQKKVLLLSEAEITFNTSVYSQLP 1092

Qy       1069 GQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1093 GQEAFLRAQTKTVLEMYKVHNPVPLIVGSSVGGLLLLAIITAILYKAGFFKRQYKEMLEE 1152
```

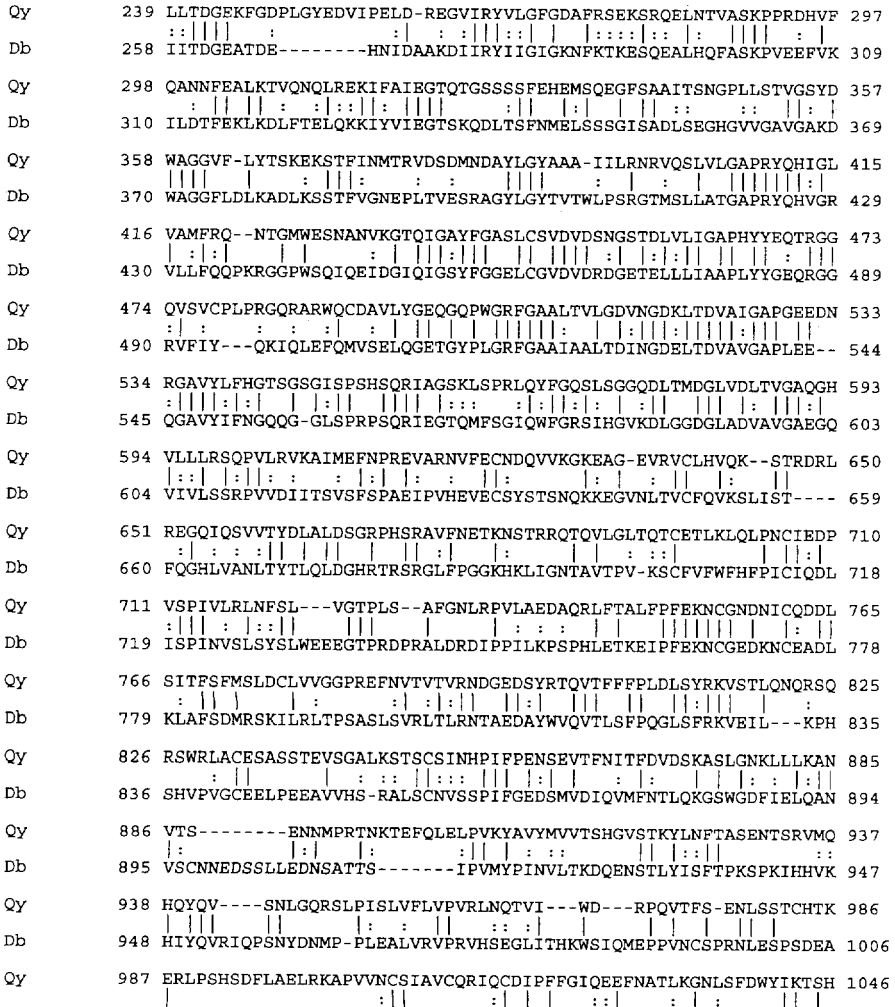
```
RESULT 10
ITAD_RAT
ID   ITAD_RAT       STANDARD;      PRT;  1161 AA.
AC   Q9QYE7;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-D precursor.
GN   Name=Itgad;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RA   O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA   Gallatin W.M.;
RT   "Cloning of rat alpha D, a novel beta 2 integrin.";
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC       VCAM1. May play a role in the atherosclerotic process such as
CC       clearing lipoproteins from plaques and in phagocytosis of blood-
CC       borne pathogens, particulate matter, and senescent erythrocytes
CC       from the blood (By similarity).
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC       associates with beta-2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF021334; AAF21241.1; -.
DR   HSSP; P11215; 1BHQ.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     19       Potential.
FT   CHAIN        20   1161       Integrin alpha-D.
FT   DOMAIN       20   1100       Extracellular (Potential).
FT   TRANSMEM   1101   1121       Potential.
FT   DOMAIN     1122   1161       Cytoplasmic (Potential).
FT   REPEAT       34     87       FG-GAP 1.
FT   REPEAT       88      ?       FG-GAP 2.
FT   DOMAIN      152    334       VWFA.
FT   REPEAT      352    402       FG-GAP 3.
FT   REPEAT      403    454       FG-GAP 4.
FT   REPEAT      456    517       FG-GAP 5.
FT   REPEAT      519    577       FG-GAP 6.
FT   REPEAT      582    634       FG-GAP 7.
FT   CA_BIND     467    475       Potential.
FT   CA_BIND     531    539       Potential.
FT   CA_BIND     594    602       Potential.
FT   SITE       1126   1130       GFFKR motif.
FT   DISULFID     69     76       By similarity.
FT   DISULFID    108    126       By similarity.
FT   DISULFID    656    711       By similarity.
FT   DISULFID    769    775       By similarity.
FT   DISULFID    845    860       By similarity.
FT   DISULFID    993   1017       By similarity.
FT   DISULFID   1022   1027       By similarity.
FT   CARBOHYD     61     61       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    245    245       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    393    393       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    696    696       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    734    734       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    784    784       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    907    907       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    936    936       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1045   1045       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   1161 AA;  126600 MW;  2258491A984A705E CRC64;
```

```
  Query Match             54.9%;  Score 3225.5;  DB 1;  Length 1161;
  Best Local Similarity   57.2%;  Pred. No. 2.1e-205;
  Matches  645;  Conservative  163;  Mismatches  306;  Indels   13;  Gaps    9;

Qy          2 NLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPIR 61
Db         21 NLDVEEPIVFREDAASFGQTVVQFGGSRLVVGAPLEAVAVNQTGRLYDCAPATGMCQPIV 80

Qy         62 LQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQKF 121
Db         81 LRSPLEAVNMSLGLSLVTATNNAQLLACGPTAQRACVKNMYAKGSCLLLGSSL-QFIQAV 139

Qy        122 PEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEFR 181
Db        140 PASMPECPRQEMDIAFLIDGSGSINQRDFAQMKDFVKALMGEFASTSTLFSLMQYSNILK 199

Qy        182 IHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLLT 241
Db        200 THFTFTEFKNILDPQSLVDPIVQLQGLTYTATGIRTVMEELFHSKNGSRKSAKKILLVIT 259

Qy        242 DGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQANN 301
Db        260 DGQKYRDPLEYSDVIPAADKAGIIRYAIGVGDAFQEPTALKELNTIGSAPPQDHVFKVGN 319

Qy        302 FEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAGG 361
Db        320 FAALRSIQRQLQEKIFAIEGTQSRSSSSFQHEMSQEGFSSALTSDGPVLGAVGSFSWSGG 379

Qy        362 VFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFRQ 421
Db        380 AFLYPPNTRPTFINMSQENVDMRDSYLGYSTAVAFWKGVHSLILGAPRHQHTGKVVIFTQ 439

Qy        422 NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPLP 481
Db        440 EARHWRPKSEVRGTQIGSYFGASLCSVDVDRDGSTDLVLIGAPHYYEQTRGGQVSVFPVP 499

Qy        482 RGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFH 541
Db        500 -GVRGRWQCEATLHGEQGHPWGRFGVALTVLGDVNGDNLADVAIGAPGEEESRGAVYIFH 558

Qy        542 GTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQP 601
Db        559 GASRLEIMPSPSQRVTGSQLSLRLQYFGQSLSGGQDLTQDGLVDLAVGAQGHVLLLRSLP 618
```

```
Qy        602 VLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVTY 661
              :|:|:  : | | |||: |::| ::     ||||  ||| | | : | |  | :|  | |
Db        619 LLKVELSIRFAPMEVAKAVYQCWERTPTVLEAGEATVCLTVHKGSPDLL--GNVQGSVRY 676

Qy        662 DLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNFS 721
              ||||| ||  |||:|:|||| |    : |||   |||:|| ||:|:|| ||||:||||||
Db        677 DLALDPGRLISRAIFDETKNCTLTGRKTLGLGDHCETVKLLLPDCVEDAVSPIILRLNFS 736

Qy        722 LVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVGG 781
              ||     |  || |||  :|   ||  ||||||  : :|: || |:|:|  |  |||||
Db        737 LVRDSASP-RNLHPVLAVGSQDHITASLPFEKNCKQELLCEGDLGISFNFSGLQVLVVGG 795

Qy        782 PREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEVS 841
              |  ||||| |:||||| | | |::|  ||||:|:  | |  |  |||||:  : :
Db        796 SPELTVTVTVWNEGEDSYGTLVKFYYPAGLSYRRVTGTQ-QPHQYPLRLACEAEPAAQED 854

Qy        842 GALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEFQ 901
              |:|:|||||||||| | :: || |||||  || ||::|||:|  :|||| | |||| ||
Db        855 --LRSSSCSINHPIFREGAKTTFMITFDVSYKAFLGDRLLLRAKASSENNKPDTNKTAFQ 912

Qy        902 LELPVKYAVYMVVTSHGVSTKYLNFTASENTSR-VMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||| || :::    || ::||::|    |    |:|:|:||    | : : | |||
Db        913 LELPVKYTVYTLISRQEDSTNHVNFSSSHGGRRQEAAHRYRVNNLSPLKLAVRVNFWVPV 972

Qy        961 RLNQTVIWDRPQVTFSENLSS-TCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              ||   :||   || |      :| :: : | : ||| ::::  |::|||| |   :|||
Db        973 LLNGVAVWD---VTLSSPAQGVSCVSQMKPPQNPDFLTQIQRRSVLDCSIADCLHFRCDI 1029

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              |   ||:| :  |:|||| |  :|    :|:|| ||| |: ||:: |||| ||:|:| |
Db       1030 PSLDIQDELDFILRGNLSFGWVSQTLQEKVLLVSEAEITFDTSVYSQLPGQEAFLRAQVE 1089

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMM 1126
              | :| : |  |: |: |||||||||||||  ||||||||||||:|:
Db       1090 TTLEEYVVYEPIFLVAGSSVGGLLLLALITVVLYKLGFFKRQYKEML 1136
```

```
RESULT 11
ITAL_BOVIN
ID   ITAL_BOVIN     STANDARD;      PRT;  1165 AA.
AC   P61625;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE   alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE   chain) (CD11a).
GN   Name=ITGAL;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=14697514; DOI=10.1016/j.gene.2003.09.043;
RA   Fett T., Zecchinon L., Baise E., Desmecht D.;
RT   "The bovine (Bos taurus) CD11a-encoding cDNA: molecular cloning,
RT   characterisation and comparison with the human and murine
RT   glycoproteins.";
RL   Gene 325:97-101(2004).
CC   -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC       ICAM3 and ICAM4. It is involved in a variety of immune phenomena
CC       including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC       mediated killing, and antibody dependent killing by granulocytes
CC       and monocytes (By similarity).
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC       associates with beta-2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY267467; AAP94035.1; -.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     23       Potential.
FT   CHAIN        24   1165       Integrin alpha-L.
FT   DOMAIN       24   1084       Extracellular (Potential).
FT   TRANSMEM   1085   1105       Potential.
FT   DOMAIN     1106   1165       Cytoplasmic (Potential).
FT   REPEAT       40     88       FG-GAP 1.
FT   REPEAT       89    146       FG-GAP 2.
FT   DOMAIN      153    324       VWFA.
FT   REPEAT      347    397       FG-GAP 3.
FT   REPEAT      398    452       FG-GAP 4.
FT   REPEAT      454    513       FG-GAP 5.
FT   REPEAT      515    572       FG-GAP 6.
FT   REPEAT      575    627       FG-GAP 7.
FT   CA_BIND     465    473       Potential.
FT   CA_BIND     527    535       Potential.
FT   CA_BIND     587    595       Potential.
FT   SITE       1111   1115       GFFKR motif.
FT   DISULFID     71     78       By similarity.
FT   DISULFID    108    126       By similarity.
FT   DISULFID    650    704       By similarity.
FT   DISULFID    768    774       By similarity.
FT   DISULFID    842    858       By similarity.
FT   DISULFID    994   1009       By similarity.
FT   DISULFID   1017   1048       By similarity.
FT   CARBOHYD     33     33       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     86     86       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    185    185       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    646    646       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    667    667       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    723    723       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    859    859       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    894    894       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    929    929       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1056   1056       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1067   1067       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   1165 AA;  128725 MW;  DAEB3A3F1E1463CB CRC64;
```

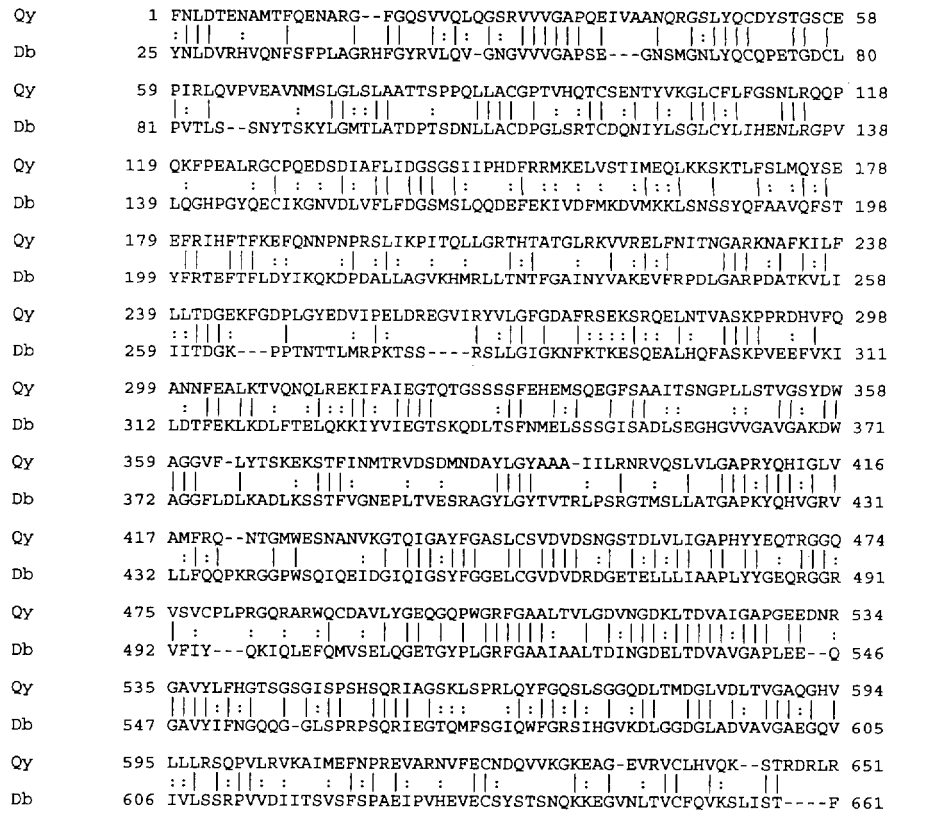
```
  Query Match             26.6%;  Score 1561;  DB 1;  Length 1165;
  Best Local Similarity   34.4%;  Pred. No. 1.4e-94;
  Matches  407;  Conservative  200;  Mismatches  468;  Indels  108;  Gaps   33;

Qy          1 FNLDTENAMTFQENARG--FGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58
              :|||  :   |      |  ||  |:|: |: |||||| |     |  |:||||   || |
Db         24 YNLDVRHVQNFSFPLAGRHFGYRVLQV-GNGVVVGAPSE---GNSMGNLYQCQPETGDCL 79

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQP 118
              |: |       : ||::||   :    |||| | : :|| :| |: |||:|   |||
Db         80 PVTLS--SNYTSKYLGMTLATDPTSDNLLACDPGLSRTCDQNIYLSGLCYLIHENLRGPV 137

Qy        119 QKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSE 178
               :    : | : : |: || ||| |:   :| :: : :  :|::|   |   |: :|:|
Db        138 LQGHPGYQECIKGNVDLVFLFDGSMSLQQDEFEKIVDFMKDVMKKLSNSSYQFAAVQFST 197

Qy        179 EFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILF 238
              ||  ||| ::    :| :|:  :  :   |:|   :  | :|:|    ||| :| |:|
Db        198 YFRTEFTFLDYIRQKDPDALLAGVKHMRLLTNTFGAINYVAKEVFRPDLGARPDATKVLI 257
```

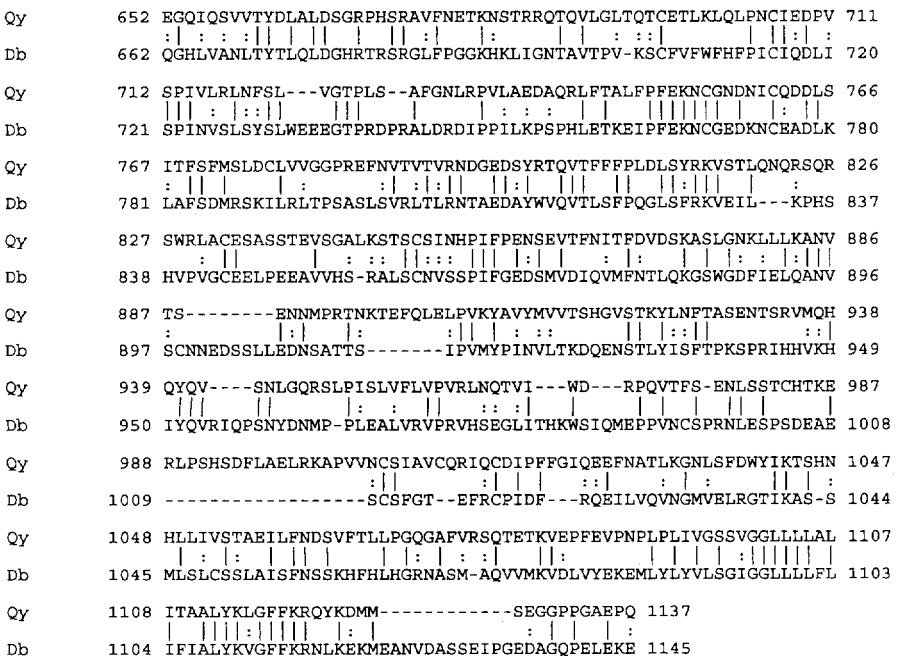


RESULT 12
Q6TYB8
ID Q6TYB8 PRELIMINARY; PRT; 1166 AA.
AC Q6TYB8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.

OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AY382558; AAQ90015.1; -.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 1.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Cell adhesion; Integrin; Transmembrane.
SQ SEQUENCE 1166 AA; 128723 MW; 7B8D8AFBA896C9DF CRC64;

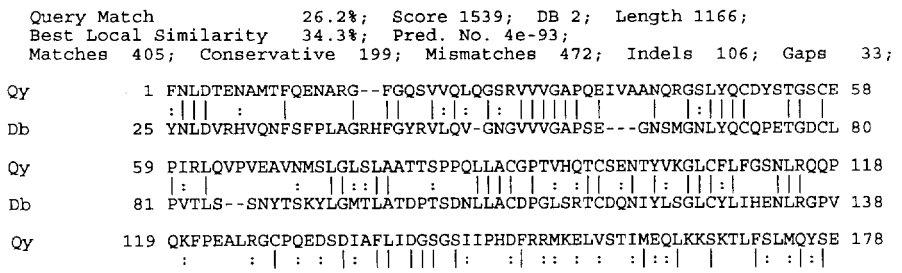
Query Match 26.2%; Score 1539; DB 2; Length 1166;
Best Local Similarity 34.3%; Pred. No. 4e-93;
Matches 405; Conservative 199; Mismatches 472; Indels 106; Gaps 33;

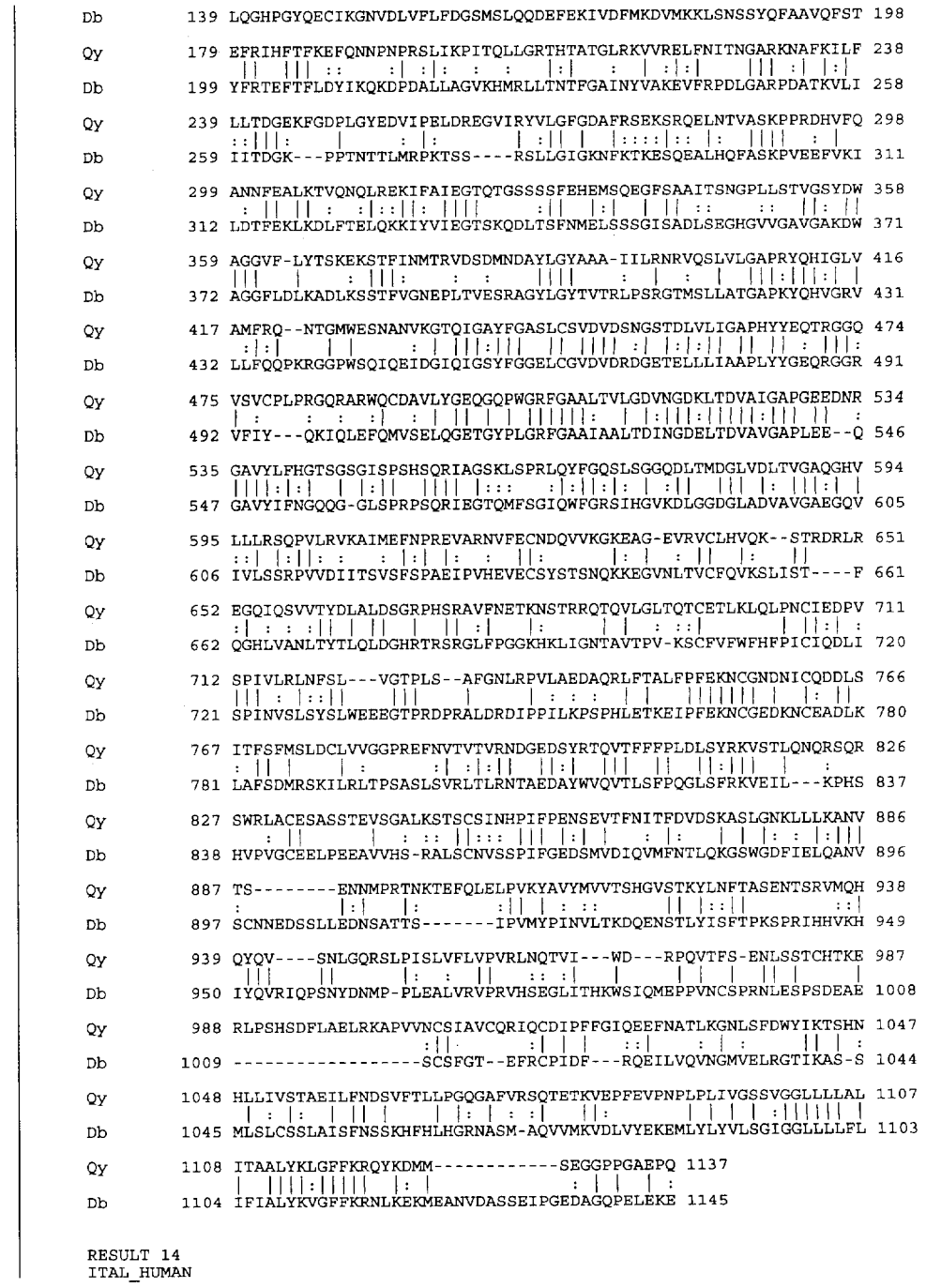




RESULT 13
AAQ90015
ID AAQ90015 PRELIMINARY; PRT; 1166 AA.
AC AAQ90015;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT "Molecular cloning and sequencing of bovine CD11a.";
RL Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY382558; AAQ90015.1; -.
SQ SEQUENCE 1166 AA; 128723 MW; 7B8D8AFBA896C9DF CRC64;

Query Match 26.2%; Score 1539; DB 2; Length 1166;
Best Local Similarity 34.3%; Pred. No. 4e-93;
Matches 405; Conservative 199; Mismatches 472; Indels 106; Gaps 33;





RESULT 14
ITAL_HUMAN

```
ID   ITAL_HUMAN     STANDARD;      PRT;  1170 AA.
AC   P20701; O43746;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE   alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE   chain) (CD11a).
GN   Name=ITGAL; Synonyms=CD11A;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RX   MEDLINE=89139587; PubMed=2537322;
RA   Larson R.S., Corbi A.L., Berman L., Springer T.;
RT   "Primary structure of the leukocyte function-associated molecule-1
RT   alpha subunit: an integrin with an embedded domain defining a protein
RT   superfamily.";
RL   J. Cell Biol. 108:703-712(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RX   MEDLINE=99425270; PubMed=10493829;
RA   Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
RA   Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA   Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
RA   Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT   "Genome duplications and other features in 12 Mb of DNA sequence from
RT   human chromosome 16p and 16q.";
RL   Genomics 60:295-308(1999).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.
RX   MEDLINE=96036067; PubMed=7479767;
RA   Qu A., Leahy D.J.;
RT   "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L
RT   beta 2) integrin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.
RX   MEDLINE=96398682; PubMed=8805579;
RA   Qu A., Leahy D.J.;
RT   "The role of the divalent cation in the structure of the I domain from
RT   the CD11a/CD18 integrin.";
RL   Structure 4:931-942(1996).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.
RX   MEDLINE=99425288; PubMed=10493852;
RA   Kallen J., Welzenbach K., Ramage P., Geyl D., Kriwacki R., Legge G.,
RA   Cottens S., Weitz-Schmidt G., Hommel U.;
RT   "Structural basis for LFA-1 inhibition upon lovastatin binding to the
RT   CD11a I-domain.";
RL   J. Mol. Biol. 292:1-9(1999).
CC   -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC       ICAM3 and ICAM4. It is involved in a variety of immune phenomena
CC       including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC       mediated killing, and antibody dependent killing by granulocytes
CC       and monocytes.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P20701-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P20701-2; Sequence=VSP_002738;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: Leukocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11a.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y00796; CAA68747.1; -.
DR   EMBL; AC002310; AAC31672.1; -.
DR   PIR; S03308; S03308.
DR   PDB; 1CQP; X-ray; A/B=153-334.
DR   PDB; 1DGQ; NMR; A=149-336.
DR   PDB; 1LFA; X-ray; A/B=-.
DR   PDB; 1MJN; X-ray; A=153-331.
DR   PDB; 1MQ8; X-ray; B/D=155-331.
DR   PDB; 1MQ9; X-ray; A=152-330.
DR   PDB; 1MQA; X-ray; A=152-330.
DR   PDB; 1ZON; X-ray; @=150-336.
DR   PDB; 1ZOO; X-ray; A/B=150-336.
DR   PDB; 1ZOP; X-ray; A/B=150-336.
DR   Genew; HGNC:6148; ITGAL.
DR   MIM; 153370; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   3D-structure; Alternative splicing; Calcium; Cell adhesion;
KW   Direct protein sequencing; Glycoprotein; Integrin; Magnesium;
KW   Receptor; Repeat; Signal; Transmembrane.
FT   SIGNAL       1     25
FT   CHAIN       26   1170       Integrin alpha-L.
FT   DOMAIN      26   1090       Extracellular (Potential).
FT   TRANSMEM  1091   1111       Potential.
FT   DOMAIN    1112   1170       Cytoplasmic (Potential).
FT   REPEAT      42     91       FG-GAP 1.
FT   REPEAT      92    149       FG-GAP 2.
FT   DOMAIN     156    327       VWFA.
FT   REPEAT     350    400       FG-GAP 3.
FT   REPEAT     401    455       FG-GAP 4.
FT   REPEAT     457    516       FG-GAP 5.
FT   REPEAT     518    575       FG-GAP 6.
FT   REPEAT     578    630       FG-GAP 7.
FT   CA_BIND    468    476       Potential.
FT   CA_BIND    530    538       Potential.
FT   CA_BIND    590    598       Potential.
FT   SITE      1115   1119       GFFKR motif.
FT   DISULFID    73     80       By similarity.
FT   DISULFID   111    129       By similarity.
FT   DISULFID   653    707       By similarity.
FT   DISULFID   771    777       By similarity.
FT   DISULFID   845    861       By similarity.
FT   DISULFID   998   1013       By similarity.
FT   DISULFID  1021   1052       By similarity.
FT   CARBOHYD    65     65       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   188    188       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   649    649       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   670    670       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   726    726       N-linked (GlcNAc. . .) (Potential).
```

```
FT   CARBOHYD    730    730       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    862    862       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    885    885       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    897    897       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1060   1060       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1071   1071       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC    954    954       Q -> QGVHGLVEMQTSKQILCRPAGDAEHTVGAQEGELPC
FT                                PWGVSEAFRDNIRAGPCR (in isoform 2).
FT                                /FTId=VSP_002738.
FT   CONFLICT    214    214       R -> W (in Ref. 1 and 2).
FT   CONFLICT    660    660       Y -> I (in Ref. 2).
FT   STRAND      155    162
FT   STRAND      164    164
FT   TURN        165    166
FT   HELIX       169    185
FT   TURN        186    188
FT   STRAND      191    198
FT   STRAND      202    206
FT   HELIX       208    214
FT   HELIX       217    221
FT   TURN        222    223
FT   STRAND      229    229
FT   HELIX       233    243
FT   TURN        244    244
FT   HELIX       247    249
FT   TURN        250    250
FT   TURN        253    254
FT   STRAND      256    263
FT   HELIX       274    276
FT   TURN        277    278
FT   STRAND      280    286
FT   HELIX       288    290
FT   HELIX       293    297
FT   TURN        298    299
FT   HELIX       300    302
FT   HELIX       307    310
FT   STRAND      311    314
FT   HELIX       317    319
FT   HELIX       323    330
SQ   SEQUENCE   1170 AA;  128819 MW;  39A7AF92EF286FC0 CRC64;

  Query Match             26.1%;  Score 1536.5;  DB 1;  Length 1170;
  Best Local Similarity   34.2%;  Pred. No. 5.8e-93;
  Matches  401;  Conservative  209;  Mismatches  463;  Indels  101;  Gaps   37;

Qy          1 FNLDTENAMTFQ--ENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58
              :|||    | :|       | ||  |:|: |: |:|||| |     |  ||||||   || |
Db         26 YNLDVRGARSFSPPRAGRHFGYRVLQV-GNGVIVGAPGE---GNSTGSLYQCQSGTGHCL 81

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLR--- 115
              |: |:     :  ||::||   :   :||| | : :|| :|||: |||:||  ||:
Db         82 PVTLR-GSNYTSKYLGMTLATDPTDGSILACDPGLSRTCDQNTYLSGLCYLFRQNLQGPM 140

Qy        116 -QQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLM 174
               |    | | ::|    : |: || ||| |: | :|::: : :  :|::|  :   |: :
Db        141 LQGRPGFQECIKG----NVDLVFLFDGSMSLQPDEFQKILDFMKDVMKKLSNTSYQFAAV 196

Qy        175 QYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAF 234
              |:|  ::   | | ::    :| :|:| :  :|  |:|   :  |  |:|    ||| :|
Db        197 QFSTSYKTEFDFSDYVKRKDPDALLKHVKHMLLLTNTFGAINYVATEVFREELGARPDAT 256

Qy        235 KILFLLTDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRD 294
              |:| ::|||       |  |   :     :|||::| |  |::::|:: |:  |||  :
Db        257 KVLIIITDGE--ATDSGNIDAAKD-----IIRYIIGIGKHFQTKESQETLHKFASKPASE 309

Qy        295 HVFQANNFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVG 354
               |   : || || :  :|::||: ||||     :||  |:|  | || ::    ::  ||
Db        310 FVKILDTFEKLKDLFTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSRGHAVVGAVG 369

Qy        355 SYDWAGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQH 412
              : |||||   |    :  |||    :  ::   ||||    :  | :   |  |||||||
Db        370 AKDWAGGFLDLKADLQDDTFIGNEPLTPEVRAGYLGYTVTWLPSRQKTSLLASGAPRYQH 429

Qy        413 IGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQT 470
              :| | :|:  |  | |      : |||||:|||  || |||| :| |:|:|||||  :| :
Db        430 MGRVLLFQEPQGGGHWSQVQTIHGTQIGSYFGGELCGVDVDQDGETELLLIGAPLFYGEQ 489

Qy        471 RGGQVSVCPLPRGQRARWQCDAV--LYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAP 528
              |||:|  :       || :   : |  | |: | | |||| |:| | |:||| | |||:|||
Db        490 RGGRVFIY-----QRRQLGFEEVSELQGDPGYPLGRFGEAITALTDINGDGLVDVAVGAP 544

Qy        529 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTV 588
              ||  :||||:|:|  | |:||  |||| |:::   :|:||:|: | :||  ||| |: |
Db        545 LEE--QGAVYIFNGRHG-GLSPQPSQRIEGTQVLSGIQWFGRSIHGVKDLEGDGLADVAV 601

Qy        589 GAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVV-KGKEAGEVRVCLHVQKSTR 647
              ||:  :::| |:||: :  :| |:| |:  :  ||:      | ||   : :|  : ||
Db        602 GAESQMIVLSSRPVVDMVTLMSFSPAEIPVHEVECSYSTSNKMKEGVNITICFQI-KSLY 660

Qy        648 DRLREGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCI 707
               :  :|:: : :|| | ||  |    | :|   ::  ||   |   : :|       | |:
Db        661 PQF-QGRLVANLTYTLQLDGHRTRRRGLFPGGRHELRRNIAVT-TSMSCTDFSFHFPVCV 718

Qy        708 EDPVSPIVLRLNFSL---VGTPLS--AFGN-----LRPVLAEDAQRLFTALFPPEKNCGN 757
              :| :||| : |||||    |||    | |      ||| | :   :      |||||||
Db        719 QDLISPINVSLNFSLWEEEGTPRDQRAQGKDIPPILRPSLHSETWEI-----PFEKNCGE 773

Qy        758 DNICQDDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVS 817
              |  |: :| ::||       | :      :| ::: |  ||:|  |:   ||:|||
Db        774 DKKCEANLRVSFSPARSRALRLTAFASLSVELSLSNLEEDAYWVQLDLHFPPGLSFRKVE 833

Qy        818 TLQNQRSQRSWRLACES--ASSTEVSGALKSTSCSINHPIFPPENSEVTFNITFDVDSKAS 875
              |  :      ::||    |  :| ||   ||::: |||      |   : |:    :|
Db        834 ML---KPHSQIPVSCEELPEESRLLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNSS 887

Qy        876 LGNKLLLKANVTSENN----MPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASEN 931
              |: : | |||| |     :   : |     :|: | : :::      || |::||
Db        888 WGDSVELHANVTCNNEDSDLLEDNSATTI---IPILYPINILIQDQEDSTLYVSFTPKGP 944

Qy        932 TSRVMQHQYQV---SNLGQRSLP-ISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTK- 986
               ::| |||    ::   ::| :  |  ||   ::  |  :  |    :   || :
Db        945 KIHQVKHMYQVRIQPSIHDHNIPTLEAVVGVPQPPSEGPITHQWSVQMEPPV--PCHYED 1002

Qy        987 -ERLPSHSD--FLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIK 1043
               ||||  ::      | : |||                    ::|    : | |      |:
Db       1003 LERLPDAAEPCLPGALFRCPVV-----------------FRQEILVQVIGTLELVGEIE 1044

Qy       1044 TSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLL 1103
              | :   : |:  | || |    | |  | : :|   ||:        | | | | :||||
Db       1045 AS-SMFSLCSSLSISFNSSKHFHLYGSNASL-AQVVMKVDVVYEKQMLYLYVLSGIGGLL 1102

Qy       1104 LLALITAALYKLGFFKRQYKDMMSEG-GPPGAEP 1136
              || ||   |||:|||||  |: |  | | |   |
Db       1103 LLLLIFIVLYKVGFFKRNLKEKMEAGRGVPNGIP 1136


RESULT 15
ITAL_MOUSE
ID   ITAL_MOUSE     STANDARD;      PRT;  1163 AA.
AC   P24063;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE   alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE   chain) (CD11a).
GN   Name=Itgal; Synonyms=Lfa-1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91268576; PubMed=2051027;
RA   Kaufmann Y., Tseng E., Springer T.A.;
RT   "Cloning of the murine lymphocyte function-associated molecule-1
```

```
RT   alpha-subunit and its expression in COS cells.";
RL   J. Immunol. 147:369-374(1991).
RN   [2]
RP   SEQUENCE OF 24-42.
RX   MEDLINE=85188276; PubMed=3887182;
RA   Springer T.A., Teplow D.B., Dreyer W.J.;
RT   "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
RT   glycoproteins and unexpected relation to leukocyte interferon.";
RL   Nature 314:540-542(1985).
CC   -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC       ICAM3 and ICAM4. Is involved in a variety of immune phenomena
CC       including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC       mediated killing, and antibody dependent killing by granulocytes
CC       and monocytes. Mice expressing a null mutation of the alpha-L
CC       subunit gene demonstrate impaired tumor rejection and impaired
CC       leukocytes recruitment.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Leukocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M60778; AAA39426.1; -.
DR   PIR; I56126; I56126.
DR   HSSP; P20701; 1DGQ.
DR   MGD; MGI:96606; Itgal.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 2.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Direct protein sequencing; Glycoprotein;
KW   Integrin; Magnesium; Receptor; Repeat; Signal; Transmembrane.
FT   SIGNAL       1     23
FT   CHAIN       24   1163       Integrin alpha-L.
FT   DOMAIN      24   1087       Extracellular (Potential).
FT   TRANSMEM  1088   1108       Potential.
FT   DOMAIN    1109   1163       Cytoplasmic (Potential).
FT   REPEAT      39     88       FG-GAP 1.
FT   REPEAT      89    146       FG-GAP 2.
FT   DOMAIN     153    325       VWFA.
FT   REPEAT     348    398       FG-GAP 3.
FT   REPEAT     399    454       FG-GAP 4.
FT   REPEAT     455    514       FG-GAP 5.
FT   REPEAT     516    573       FG-GAP 6.
FT   REPEAT     576    628       FG-GAP 7.
FT   CA_BIND    466    474       Potential.
FT   CA_BIND    528    536       Potential.
FT   CA_BIND    588    596       Potential.
FT   SITE      1111   1115       GFFKR motif.
FT   DISULFID     70     77       By similarity.
FT   DISULFID    108    126       By similarity.
FT   DISULFID    147    199       By similarity.
FT   DISULFID    651    705       By similarity.
FT   DISULFID    767    773       By similarity.
FT   DISULFID    840    856       By similarity.
FT   DISULFID    993   1009       By similarity.
FT   DISULFID   1017   1048       By similarity.
FT   CARBOHYD     86     86       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    185    185       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    270    270       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    444    444       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    668    668       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    696    696       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    724    724       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    728    728       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    776    776       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    857    857       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    880    880       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    890    890       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    899    899       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    927    927       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1056   1056       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   1163 AA;  128343 MW;  A7A3078489E8232F CRC64;
```

Query Match 25.8%; Score 1516.5; DB 1; Length 1163;
Best Local Similarity 33.9%; Pred. No. 1.2e-91;
Matches 396; Conservative 216; Mismatches 466; Indels 89; Gaps 36;

```
Qy          1 FNLDTENAMTFQENA-RGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEP 59
              :|||     :|   | | ||  |:|::   ||||||  |     :  | || |  |:  |:|
Db         24 YNLDTRPTQSFLAQAGRHFGYQVLQIEDG-VVVGAPGE---GDNTGGLYHCRTSSEFCQP 79

Qy         60 IRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQ 119
              : |          :  ||::||   :   ||||  | : :|| :|||: |||:||  :|
Db         80 VSLH-GSNHTSKYLGMTLATDAAKGSLLACDPGLSRTCDQNTYLSGLCYLFPQSLEGPML 138

Qy        120 KFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEE 179
              :   | : | :   |: || ||| |:   || :: | :  :| :|  :   |: :|:| :
Db        139 QNRPAYQECMKGKVDLVFLFDGSQSLDRKDFEKILEFMKDVMRKLSNTSYQFAAVQFSTD 198

Qy        180 FRIHFTFKEF-QNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILF 238
              |  ||| :: : | ||  |:  :  :   |:|   :  ||  :|   :||| :| |:|
Db        199 CRTEFTFLDYVKQNKNPDVLLGSVQPMFLLTNTFRAINYVVAHVFKEESGARPDATKVLV 258

Qy        239 LLTDGEKF--GDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHV 296
              ::||||    |:      |          : ||::| |  | | : :: |:  ||:|  : |
Db        259 IITDGEASDKGNISAAHD---------ITRYIIGIGKHFVSVQKQKTLHIFASEPVEEFV 309

Qy        297 FQANNFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSY 356
                 : || || :   |: :|:|||||      :||  |:|  | || ::    ::  ||:
Db        310 KILDTFEKLKDLFTDLQRRIYAIEGTNRQDLTSFNMELSSSGISADLSKGHAVVGAVGAK 369

Qy        357 DWAGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQHIG 414
              |||||   |    : :||:    : ||:   ||||  | :  |:    |  |||||||:|
Db        370 DWAGGFLDLREDLQGATFVGQEPLTSDVRGGYLGYTVAWMTSRSSRPLLAAGAPRYQHVG 429

Qy        415 LVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRG 472
              | :|:  :  | |      ::|||||:|||  |||||:| :|  :|:|||||  :: : ||
Db        430 QVLLFQAPEAGGRWNQTQKIEGTQIGSYFGGELCSVDLDQDGEAELLLIGAPLFFGEQRG 489

Qy        473 GQVSVCPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEED 532
              |:|      : ::: ::  : | |:  | |  ||||||:| | |:|||:|||||:||| ||
Db        490 GRVFTY---QRRQSLFEMVSELQGDPGYPLGRFGAAITALTDINGDRLTDVAVGAPLEE- 545

Qy        533 NRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQG 592
              :||||:|:|  | |:||  |||| |::: | :::||:|: | :||  | | |: |||:|
Db        546 -QGAVYIFNGKPG-GLSPQPSQRIQGAQVFPGIRWFGRSIHGVKDLGGDRLADVVVGAEG 603

Qy        593 HVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAG-EVRVCLHVQKSTRDRLR 651
              |::| |:||: |   : |:| |:  :  ||:    : :: | ::: |  ::  |
Db        604 RVVVLSSRPVVDVVTELSFSPEEIPVHEVECSYSAREEQKHGVKLKACFRIKPLTPQ--F 661

Qy        652 EGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPV 711
              :|:: : ::| | ||  |  || :| :  :     | :    ::|       | ||:| :
Db        662 QGRLLANLSYTLQLDGHRMRSRGLFPDGSHELSGNTSITP-DKSCLDFHFHFPICIQDLI 720

Qy        712 SPIVLRLNFSLV---GTPLSAFGN-LRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSI 767
```

```
Db        721 SPINVSLNFSLLEEEGTPRDQKGRAMQPILRPSIHTV-TKEIPFEKNCGEDKKCEANLTL 779

Qy        768 TFSFMSLDCLVVGGP------REFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQN 821

Db        780 SSPARS-------GPLRLMSSASLAVEWTLSNSGEDAYWVRLDLDFPRGLSFRKVEMLQ- 831

Qy        822 QRSQRSWRLACESASSTEVSGAL-KSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKL 880

Db        832 --PHSRMPVSCEEL--TEGSSLLTKTLKCNVSSPIFKAGQEVSLQVMFNTLLNSSWEDFV 887

Qy        881 LLKANVTSEN-NMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQ 939

Db        888 ELNGTVHCENENSSLQEDNSAATHIPVLYPVNILTKEQENSTLYISFTPKGPKTQQVQHV 947

Qy        940 YQVSNLGQRSLPISLVFLVPVRLNQTVIWDRPQ----VTFSENLSS----TCHTKE-RLP 990

Db        948 YQV-----RIQPSAYDHNMPT-LEALVGVPRPHSEDLITYTWSVQTDPLVTCHSEDLKRP 1001

Qy        991 SHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLL 1050

Db       1002 SSE---AEQPCLPGV--------QFRCPIVF---RWEILIQVTGTVELSKEIKAS-STLS 1046

Qy       1051 IVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLALITA 1110

Db       1047 LCSSLSVSFNSSKHFHLYGSKA-SEAQVLVKVDLIHEKEMLHVYVLSGIGGLVLLFLIFL 1105

Qy       1111 ALYKLGFFKRQYKDMM-SEGGPPGAEP 1136

Db       1106 ALYKVGFFKRNLKEKMEADGGVPNGSP 1132

Search completed: November 9, 2004, 12:17:27
Job time : 142.75 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:45:42 ; Search time 103.25 Seconds
(without alignments)
3950.365 Million cell updates/sec

Title: US-09-902-481B-3
Perfect score: 5879
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5839	99.3	1153	2	AAW65090	Aaw65090 Human Bet
2	5839	99.3	1153	3	AAB07360	Aab07360 Human CD1
3	5839	99.3	1153	5	AAU80252	Aau80252 Human int
4	5839	99.3	1153	5	ABG61469	Abg61469 Human Bet
5	5839	99.3	1153	5	AAO14428	Aao14428 Integrin
6	5839	99.3	1153	7	ADD25615	Add25615 Binding d
7	5829	99.1	1153	2	AAR04136	Aar04136 Alpha sub
8	5823.5	99.1	1152	8	ADM99589	Adm99589 Human int
9	5816.5	98.9	1152	8	ADP12435	Adp12435 Protein e
10	3460	58.9	1163	8	ADP44061	Adp44061 Human CD1
11	3446	58.6	1163	8	ADN02004	Adn02004 Human inf
12	3446	58.6	1163	8	ADQ17510	Adq17510 Human sof
13	3437	58.5	1163	2	AAR07120	Aar07120 p150.95 a
14	3423	58.2	1163	2	AAW65091	Aaw65091 Human Bet
15	3423	58.2	1163	3	AAB07361	Aab07361 Human CD1
16	3423	58.2	1163	5	ABG61470	Abg61470 Human Bet
17	3421	58.2	1163	6	ABU07406	Abu07406 Protein d
18	3421	58.2	1163	7	ADG32005	Adg32005 Human hom
19	3388	57.6	1161	2	AAR78166	Aar78166 Human bet
20	3388	57.6	1161	2	AAW23049	Aaw23049 Human bet
21	3388	57.6	1161	2	AAW57491	Aaw57491 Human bet
22	3388	57.6	1161	2	AAW65089	Aaw65089 Human Bet
23	3388	57.6	1161	2	AAW72825	Aaw72825 Human alp
24	3388	57.6	1161	2	AAW73342	Aaw73342 Human alp
25	3388	57.6	1161	3	AAB07359	Aab07359 Human alp
26	3388	57.6	1161	5	ABG61468	Abg61468 Human Bet
27	3372.5	57.4	1161	2	AAW23064	Aaw23064 Human bet
28	3372.5	57.4	1161	2	AAW65106	Aaw65106 Human Bet
29	3372.5	57.4	1161	2	AAW72837	Aaw72837 Human alp
30	3372.5	57.4	1161	2	AAW73343	Aaw73343 Human alp
31	3372.5	57.4	1161	3	AAB07376	Aab07376 Human alp
32	3372.5	57.4	1161	5	ABG61485	Abg61485 Human Bet
33	3215.5	54.7	1161	2	AAR78169	Aar78169 Rat alpha
34	3213.5	54.7	1161	2	AAW23062	Aaw23062 Rat beta
35	3213.5	54.7	1161	2	AAW60004	Aaw60004 Rat alpha
36	3213.5	54.7	1161	2	AAW72824	Aaw72824 Rat alpha
37	3213.5	54.7	1161	3	AAB07374	Aab07374 Rat alpha
38	3213.5	54.7	1161	5	ABG61483	Abg61483 Rat Beta2
39	3206.5	54.5	1161	2	AAW65104	Aaw65104 Rat beta-
40	3206.5	54.5	1161	2	AAW73345	Aaw73345 Rat alpha
41	3201	54.4	1161	2	AAW23061	Aaw23061 Mouse bet
42	3201	54.4	1161	2	AAW60003	Aaw60003 Mouse alp
43	3201	54.4	1161	2	AAW65103	Aaw65103 Mouse bet
44	3201	54.4	1161	2	AAW72836	Aaw72836 Mouse alp
45	3201	54.4	1161	2	AAW73347	Aaw73347 Mouse alp

ALIGNMENTS

RESULT 1
AAW65090
ID AAW65090 standard; protein; 1153 AA.
XX
AC AAW65090;
XX
DT 28-SEP-1998 (first entry)
XX
DE Human Beta-integrin CD11b subunit protein.
XX
KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN US5728533-A.
XX
PD 17-MAR-1998.
XX
PF 07-JUN-1995; 95US-00485618.
XX
PR 23-DEC-1993; 93US-00173497.
PR 05-AUG-1994; 94US-00286889.
PR 21-DEC-1994; 94US-00362652.
XX
PA (ICOS-) ICOS CORP.
XX
PI Van Der Vieren M, Gallatin WM;
XX
DR WPI; 1998-206565/18.
XX
PT Screening assay for modulators of integrin binding - using immobilised or
PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.
XX
PS Example 5; Fig 1A-D; 106pp; English.
XX
CC This sequence represents a human beta-integrin CD11b subunit which is
CC used to describe a method for identifying compounds that modulate the
CC interaction of the beta-integrin alpha-d subunit with a binding partner
CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha
CC -d binding partner, one of which is immobilised and the other of which is
CC labelled, in the presence of a test compound, and determining if the
CC compound affects binding between the alpha-d polypeptide and alpha-d
CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment
CC comprising the cytoplasmic, transmembrane or extracellular domain of
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

```
CC   diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC   asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC   and rheumatoid arthritis
XX
SQ   Sequence 1153 AA;

  Query Match             99.3%;  Score 5839;  DB 2;  Length 1153;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||| |||:|||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||| ||||||||:| ||||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 2
AAB07360
ID   AAB07360 standard; protein; 1153 AA.
XX
AC   AAB07360;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Human CD11b protein sequence.
XX
KW   Human; macrophage infiltration inhibition; alpha_d integrin;
KW   leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW   inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW   atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW   lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW   rheumatoid arthritis; central nervous system injury; CD11b.
XX
OS   Homo sapiens.
XX
PN   WO200029446-A1.
XX
PD   25-MAY-2000.
XX
PF   16-NOV-1999;   99WO-US027139.
XX
PR   16-NOV-1998;   98US-00193043.
PR   08-JUL-1999;   99US-00350259.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin MW,  Van Der Vieren M;
XX
DR   WPI; 2000-387751/33.
XX
PT   Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT   macrophage infiltration and reduce inflammation at central nervous system
PT   injury sites.
XX
PS   Example 5; Fig 1; 270pp; English.
XX
CC   Integrins are a class of membrane-associated molecules that participate
CC   in cellular adhesion. Integrins are made up of an alpha subunit and a
CC   beta subunit. One class of human integrins are restricted to expression
CC   in white blood cells and have a common beta2 subunit: the leukocyte
CC   integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC   have an important role in immune and inflammatory responses. The present
CC   protein sequence is the human integrin alpha subunit CD11b. This sequence
CC   was used in an alignment to identify a novel beta2 integrin alpha
CC   subunit: alpha_d (AAA60014 and AAB07359). The present sequence has
CC   approximately 60% identity to the protein sequence of alpha_d. The
CC   Alpha_d gene and protein may be useful in therapy for diseases linked to
CC   aberrant alpha_d function e.g. Type I diabetes, atherosclerosis, multiple
CC   sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
CC   distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
CC   (LAD). In addition, anti-alpha_d monoclonal antibodies may be used in the
CC   inhibition of macrophage infiltration at the site of a central nervous
CC   system injury. The monoclonal antibodies can also be used to detect and
CC   diagnose Crohn's disease
XX
```

SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 3; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:|||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||| |||||||||:| ||||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 3
AAU80252
ID AAU80252 standard; protein; 1153 AA.
XX
AC AAU80252;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human integrin 1 alpha-M subunit protein.
XX
KW Integrin; antiinflammatory; imunosuppression; nephritis; dermatitis;
KW inflammatory disease; autoimmune disorder; Crohn's disease;
KW human immunodeficiency virus; HIV; myocardial infarction;
KW Sjorgen's syndrome; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 499. .500
FT /note= "Encoded by GGG CAG AGG"
XX
PN WO200218583-A2.
XX
PD 07-MAR-2002.
XX
PF 31-AUG-2001; 2001WO-US027227.
XX
PR 01-SEP-2000; 2000US-0229700P.
XX
PA (BLOO-) CENT BLOOD RES INC.
XX
PI Springer TA, Shimoaka M, Lu C;
XX
DR WPI; 2002-382964/41.
DR N-PSDB; ABK50046.
XX
PT Modified integrin-I or integrin I-like domain polypeptide useful as an
PT immunogen to produce antibodies specific to polypeptide, comprises a
PT disulfide bond such that polypeptide is stabilized in a desired
PT conformation.
XX
PS Disclosure; Page 109-112; 112pp; English.
XX
CC This invention relates to a modified integrin-I or integrin I-like domain
CC polypeptide comprising at least one disulfide bond so that the domain is
CC stabilised in a desired conformation. The polypeptide of the invention
CC may have antiinflammatory or imunosuppressive activities. The
CC polypeptides of the invention have an open conformation and are useful as
CC immunogens to produce antibodies that selectively bind to integrin I-
CC domain; and for identifying a modulator of integrin activity, or of
CC interaction of an integrin and a cognate ligand. The polypeptide of the
CC invention, or antibodies (preferably anti-LFA-1 antibody) is useful for
CC treating or preventing an integrin mediated disorder which is an
CC inflammatory or autoimmune disorder in a subject and for inhibiting the
CC binding of an integrin to a cognate ligand such as Crohn's disease,
CC nephritis; human immunodeficiency virus (HIV), myocardial infarction,
CC Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic
CC composition comprising the peptide of the invention is useful for
CC treating an integrin mediated disorder in a subject. The polypeptides
CC and/or active or antigenic fragments are useful as reagents for diagnosis
CC of integrin-mediated disorders. The present sequence represents the human
CC integrin-1 alpha-M protein subunit used to generate the mutant
CC polypeptides of the invention

```
XX
SQ   Sequence 1153 AA;

  Query Match             99.3%;  Score 5839;  DB 5;  Length 1153;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||||| |||:|||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||| |||||||||:| ||||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 4
ABG61469
ID   ABG61469 standard; protein; 1153 AA.
XX
AC   ABG61469;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human Beta2 integrin alphaCD11b subunit.
XX
KW   Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; LAD;
KW   leukocyte adhesion deficiency; inflammatory response; diabetes;
KW   multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW   inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW   immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
KW   intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW   locomotor recovery; locomotor damage; locomotor impairment;
KW   autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX
OS   Homo sapiens.
XX
PN   WO200230980-A2.
XX
PD   18-APR-2002.
XX
PF   15-OCT-2001; 2001WO-US032059.
XX
PR   13-OCT-2000; 2000US-00688307.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin WM,  Van Der Vieren M;
XX
DR   WPI; 2002-463260/49.
XX
PT   Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT   recovery, inhibiting locomotor damage, limiting locomotor impairment, or
PT   limiting autonomic and sensory dysfunction following spinal cord injury.
XX
PS   Example 5; Page 191-194; 270pp; English.
XX
CC   The invention relates to promoting locomotor recovery, inhibiting
CC   locomotor damage, limiting locomotor impairment, or limiting autonomic
CC   and sensory dysfunction following spinal cord injury by administering an
CC   anti-alpha_d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
CC   spinal cord injury victim. The method also involves the use of a ligand
CC   selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,
CC   vascular cell adhesion molecule). The method is useful for promoting
CC   locomotor recovery, inhibiting locomotor damage, limiting locomotor
CC   impairment, or limiting autonomic and sensory dysfunction following
CC   spinal cord injury. In particular, the spinal cord injury comprises
CC   compression of the spinal cord. The antibodies are also useful for
CC   reducing inflammation at the site of a central nervous system injury. The
CC   specification also details the identification of Beta2 integrin alphaD
CC   cDNAs and proteins, for use in raising the antibodies. Beta2 integrins
CC   are implicated in diseases such as LAD (leukocyte adhesion deficiency,
CC   inflammatory response, diabetes, multiple sclerosis, arthritis, graft
CC   atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative
CC   colitis, immune complex alveolitis and leukaemia. The present sequence is
CC   a Beta2 integrin alpha subunit sequence included for comparison with the
CC   Beta2 integrin alphaD protein sequences
XX
```

SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 5; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:|||||||||||||:|||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||| |||||||||:| ||||||||||||||:|||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036
```

```
Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 5
AAO14428
ID AAO14428 standard; protein; 1153 AA.
XX
AC AAO14428;
XX
DT 03-MAY-2002 (first entry)
XX
DE Integrin Mac-1 alpha subunit.
XX
KW Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
KW open conformation; integrin related inflammatory disorder;
KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;
KW reperfusion; hypovolemic shock; infarction; cerebral shock;
KW viral infection; cancer; gene therapy; vaccine;
KW bioactive agent screening.
XX
OS Unidentified.
XX
PN WO200204521-A2.
XX
PD 17-JAN-2002.
XX
PF 09-JUL-2001; 2001WO-US021805.
XX
PR 07-JUL-2000; 2000US-0216600P.
XX
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
PA (BLOO-) CENT BLOOD RES.
XX
PI Springer T;
XX
DR WPI; 2002-148167/19.
XX
PT New integrin I domain protein having alteration in at least 2
PT noncontiguous regions and exits in an open conformation, useful for
PT treating, preventing or suppressing inflammatory or immunological
PT disorders.
XX
PS Example 1; Fig 1F; 90pp; English.
XX
CC The invention comprises structurally biased variant integrin inserted (I)
CC domain proteins, wherein the alterations to the protein occur in at least
CC two noncontinuous regions. Specifically the the variant integrin I domain
CC proteins are structurally biased to exist in the open conformation,
CC thereby altering the binding ability of the protein. The invention also
CC comprises nucleic acids encoding the variant integrin I domain proteins.
CC The integrin I domain proteins and nucleic acids are useful for treating,
CC preventing or suppressing integrin related inflammatory and immunological
CC disorders (e.g. rheumatoid arthritis). The variant integrin I domain
CC proteins and nucleic acids can also be used for treating: ischaemia/
CC reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral
CC infection; and cancer. The variant integrin I domain nucleic acids and
CC proteins may be used in gene therapy, as vaccines and to screen for
CC bioactive agents. The present amino acid sequence represents the Mac-1
CC alpha subunit of integrin
XX
SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 5; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:|||||||||||||:|||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||| |||||||||:| ||||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX
AC ADD25615;
XX
DT 15-JAN-2004 (first entry)
XX
DE Binding domain-immunoglobulin fusion protein-associated protein #85.
XX
KW Binding domain; immunoglobulin; fusion protein; cytostatic;
KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;
KW neuroprotective; hinge region; immunoglobulin heavy chain;
KW CH2 constant region; CH3 constant region; IgG1;
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.
XX
OS Unidentified.
XX
PN US2003118592-A1.
XX
PD 26-JUN-2003.
XX
PF 25-JUL-2002; 2002US-00207655.
XX
PR 17-JAN-2001; 2001US-0367358P.
PR 17-JAN-2002; 2002US-00053530.
PR 03-JUN-2002; 2002US-0385691P.
XX
PA (GENE-) GENECRAFT INC.
XX
PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
XX
DR WPI; 2003-801317/75.
XX
PT New binding domain-immunoglobulin fusion protein, useful for treating a
PT subject having or suspected of having a malignant condition or a B-cell
PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX
PS Disclosure; SEQ ID NO 176; 157pp; English.
XX
CC The invention relates to a binding domain-immunoglobulin fusion protein
CC comprising a binding domain polypeptide that is fused to an
CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
CC CH2 constant region polypeptide that is fused to the hinge region
CC polypeptide, and an immunoglobulin heavy chain CH3 constant region
CC polypeptide that is fused to the CH2 constant region polypeptide. The
CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin
CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge
CC region polypeptide, derived from (a) having 3 or more cysteine residues;
CC where the mutated human IgG1 immunoglobulin hinge region polypeptide
CC contains 2 cysteine residues, where the first cysteine is not mutated; a
CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
CC (a) having 3 or more cysteine residues, where the mutated human IgG1
CC immunoglobulin hinge region polypeptide contains no more than one
CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
CC polypeptide, derived from (a) having 3 or more cysteine residues; where
CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains
CC no cysteine residues. The binding domain-immunoglobulin fusion protein is
CC capable of at least one immunological activity comprising antibody
CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The
CC binding domain polypeptide is capable of specifically binding to an
CC antigen. Also included are an isolated polynucleotide encoding the
CC binding domain-immunoglobulin fusion protein, a recombinant expression
CC construct comprising the polynucleotide (operably linked to a promoter),
CC a host cell transformed or transfected with a recombinant expression
CC construct, producing the binding domain-immunoglobulin fusion protein, a

```
CC   pharmaceutical composition comprising the binding domain-immunoglobulin
CC   fusion protein or polynucleotide and a carrier, and treating a subject
CC   having or suspected of having a malignant condition or a B-cell disorder.
CC   The binding domain-immunoglobulin fusion protein is useful for treating a
CC   subject having or suspected of having a malignant condition or a B-cell
CC   disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,
CC   myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple
CC   sclerosis or autoimmune disease. The present sequence is a binding domain
CC   -immunoglobulin fusion protein-associated protein sequence. Note: The
CC   sequence data for this patent formed part of the printed specification
CC   and is also available in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not
CC   identified the sequences in the printed specification by their SEQ ID
CC   number therefore none of the sequences can be explicitly identified.
XX
SQ   Sequence 1153 AA;

  Query Match             99.3%;  Score 5839;  DB 7;  Length 1153;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||:|||||||||||||:|||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||| |||||||||:| ||||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 7
AAR04136
ID   AAR04136 standard; protein; 1153 AA.
XX
AC   AAR04136;
XX
DT   09-SEP-2004  (revised)
DT   25-MAR-2003  (revised)
DT   07-SEP-1990  (first entry)
XX
DE   Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX
KW   Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
KW   non-specific defence system; integrin gene superfamily.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .16
FT                   /label= signal_peptide
FT   Modified-site   86. .88
FT                   /label= putative N-glycosylation site
FT   Modified-site   240. .242
FT                   /label= putative N-glycosylation site
FT   Modified-site   391. .393
FT                   /label= putative N-glycosylation site
FT   Modified-site   469. .471
FT                   /label= putative N-glycosylation site
FT   Modified-site   693. .695
FT                   /label= putative N-glycosylation site
FT   Modified-site   697. .699
FT                   /label= putative N-glycosylation site
FT   Modified-site   735. .737
FT                   /label= putative N-glycosylation site
FT   Modified-site   802. .804
FT                   /label= putative N-glycosylation site
FT   Modified-site   881. .883
FT                   /label= putative N-glycosylation site
FT   Modified-site   901. .903
FT                   /label= putative N-glycosylation site
FT   Modified-site   912. .914
FT                   /label= putative N-glycosylation site
FT   Modified-site   941. .943
FT                   /label= putative N-glycosylation site
FT   Modified-site   947. .949
FT                   /label= putative N-glycosylation site
FT   Modified-site   979. .981
FT                   /label= putative N-glycosylation site
```

```
FT   Modified-site   994. .996
FT                   /label= putative N-glycosylation site
FT   Modified-site   1022. .1024
FT                   /label= putative N-glycosylation site
FT   Modified-site   1045. .1047
FT                   /label= putative N-glycosylation site
FT   Modified-site   1051. .1053
FT                   /label= putative N-glycosylation site
FT   Modified-site   1076. .1078
FT                   /label= putative N-glycosylation site
FT   Region          1106. .1134
FT                   /label= putative_transmembrane_region
XX
PN   EP364690-A.
XX
PD   25-APR-1990.
XX
PF   17-AUG-1989;   89EP-00115159.
XX
PR   23-AUG-1988;   88US-00235353.
PR   09-MAR-1989;   89US-00321239.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Springer TA,  Corbi A;
XX
DR   WPI; 1990-125938/17.
DR   N-PSDB; AAQ04043.
XX
PT   New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating
PT   inflammation and viral infections, and in diagnosis.
XX
PS   Disclosure; Page ?; 3pp; English.
XX
CC   Mac-1 alpha subunit is involved in the response to inflammation, i.e.
CC   recognition of and migration to sites of inflammation. It also attaches
CC   to cellular substrates as part of this function making it useful in
CC   visualising endothelial tissue. Mac-1 is a member of the Integrin Gene
CC   superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25
CC   -MAR-2003 to correct PA field.)
CC
CC   Revised record issued on 09-SEP-2004 : Correction to feature table key
CC   and pages
XX
SQ   Sequence 1153 AA;

  Query Match             99.1%;  Score 5829;  DB 2;  Length 1153;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 1123;  Conservative    8;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||| |||:||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||:|||||||||||||||:||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||| ||||||||:| ||||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPEVPNPLLLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 8
ADM99589
ID   ADM99589 standard; protein; 1152 AA.
XX
AC   ADM99589;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Human integrin alphaM subunit precursor protein.
XX
KW   integrin alpha subunit; beta; antipsoriatic; thrombolytic; anticoagulant;
KW   osteopathic; cytostatic; immunosuppressive; antiinflammatory;
KW   neuroprotective; antisickling; immunotherapy; inflammatory;
KW   autoimmune disorder; thrombosis; cancer; osteoporosis;
KW   sickle cell anaemia; psoriasis; multiple sclerosis; human; precursor;
KW   alphaM.
XX
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   Misc-difference 965
FT                   /note= "Encoded by CCC"
XX
PN   WO2004007530-A2.
XX
PD   22-JAN-2004.
XX
PF   17-JUL-2003; 2003WO-US022301.
XX
PR   17-JUL-2002; 2002US-0396783P.
PR   17-JUL-2002; 2002US-0396790P.
PR   11-SEP-2002; 2002US-0410135P.
XX
PA   (BLOO-) CENT BLOOD RES INC.
XX
PI   Springer TA,  Takagi J;
XX
DR   WPI; 2004-122877/12.
DR   N-PSDB; ADM99588.
XX
PT   Novel modified integrin protein having extracellular domains of integrin
PT   alpha and beta subunits or integrin alphaII and beta3 subunit, useful for
PT   treating integrin mediated disorders.
XX
PS   Disclosure; SEQ ID NO 4; 232pp; English.
XX
CC   The invention relates to a novel isolated or recombinant modified
CC   integrin protein having extracellular domains of integrin alpha and beta
CC   subunits where one of the subunits has one or more mutations, an altered
CC   surface feature or an amino acid substitution or internal deletion,
CC   extracellular domains of the integrin beta subunit that comprise a
CC   mutation that alters a non-cysteine residue to cysteine or extracellular
CC   domains of integrin alpha and beta subunits. The polypeptide of the
CC   invention demonstrates antipsoriatic, thrombolytic, anticoagulant,
CC   osteopathic, cytostatic, immunosuppressive, antiinflammatory,
CC   neuroprotective and antisickling activities and may be useful for
CC   immunotherapy in order to prevent or treat an integrin-mediated disorder
CC   such as an inflammatory disorder, an autoimmune disorder, thrombosis,
CC   cancer, osteoporosis, sickle cell anaemia, psoriasis and multiple
CC   sclerosis. The current sequence is that of the human integrin alphaM
CC   subunit precursor protein of the invention.
XX
SQ   Sequence 1152 AA;

  Query Match             99.1%;  Score 5823.5;  DB 8;  Length 1152;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 1123;  Conservative    8;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||||| |||:||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||:|||||||||||||:|||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||| |||||||||:| ||||||||||||||:|||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 9
ADP12435
ID   ADP12435 standard; protein; 1152 AA.
XX
AC   ADP12435;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Protein encoded by mRNA of the invention #45.
XX
KW   transplant rejection; immune system; rheumatoid arthritis; lupus;
KW   inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX
OS   Homo sapiens.
XX
PN   WO2004042346-A2.
XX
PD   21-MAY-2004.
XX
PF   24-APR-2003; 2003WO-US012946.
```

```
XX
PR   24-APR-2002; 2002US-00131831.
PR   20-DEC-2002; 2002US-00325899.
XX
PA   (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI   Wohlgemuth J,  Fry K,  Woodward R,  Ly N,  Prentice J,  Morris M;
PI   Rosenberg S;
XX
DR   WPI; 2004-400724/37.
XX
PT   Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT   pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT   rejection, in an individual, comprises detecting the expression level of
PT   the genes.
XX
PS   Claim 65; SEQ ID NO 2444; 1762pp; English.
XX
CC   The present invention relates to diagnosing or monitoring transplant
CC   rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC   comprises detecting the expression level of one or more genes. The
CC   methods, system and kits are useful in diagnosing or monitoring
CC   transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC   islet, lung, bone marrow or stem cell transplant rejection,
CC   xenotransplant rejection or mechanical organ replacement rejection, in an
CC   individual. The method is also useful in assessing the immune status of
CC   an individual. The methods are also useful in diagnosing and monitoring
CC   diseases that involve the immune system, e.g. rheumatoid arthritis,
CC   lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC   viral, bacterial or fungal infection. The present sequence represents a
CC   protein that is encoded by the mRNA of the invention.
XX
SQ   Sequence 1152 AA;

  Query Match             98.9%;  Score 5816.5;  DB 8;  Length 1152;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1122;  Conservative    8;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||| |||||||||:| ||||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 10
ADP44061
ID   ADP44061 standard; protein; 1163 AA.
XX
AC   ADP44061;
XX
DT   09-SEP-2004  (first entry)
XX
DE   Human CD11C protein SEQ ID NO:14.
XX
KW   HIV entry inhibitor; cell surface protein inhibitor; HIV infection;
KW   anti-HIV; virucide; HIV uptake inhibitor; human; CD11C; integrin alpha X.
XX
OS   Homo sapiens.
XX
PN   WO2004053094-A2.
XX
PD   24-JUN-2004.
XX
PF   08-DEC-2003; 2003WO-US039208.
XX
PR   06-DEC-2002; 2002US-0431522P.
XX
PA   (PPDP-) PPD DEV LP.
XX
PI   Dunn SJ,  Holzmayer TA;
XX
DR   WPI; 2004-480928/45.
DR   N-PSDB; ADP44060.
XX
PT   Identifying an inhibitor of human immunodeficiency virus (HIV) entry into
```

PT a human host cell useful for preventing and/or treating HIV infection, by
PT identifying an inhibitor of a cell surface polypeptide such as CXCR-4.
XX
PS Claim 1; SEQ ID NO 14; 133pp; English.
XX
CC The present invention describes a method (M1) for identifying an
CC inhibitor of HIV entry into a human host cell. (M1) comprises identifying
CC an inhibitor of a cell surface polypeptide selected from CXCR-4 (352
CC amino acids, SEQ ID NO:8, ADP44055), CCR4 (360 amino acids, SEQ ID NO:10,
CC ADP44057), CCR7 (378 amino acids, SEQ ID NO:12, ADP44059), CD11C (1163
CC amino acids, SEQ ID NO:14, ADP44061), CD47 (323 amino acids, SEQ ID
CC NO:16, ADP44063), CD68 (354 amino acids, SEQ ID NO:18, ADP44065), CD69
CC (199 amino acids, SEQ ID NO:20, ADP44067), CD74 (566 amino acids, SEQ ID
CC NO:22, ADP44069), CSF3R (836 amino acids, SEQ ID NO:24, ADP44071), RARA
CC (462 amino acids, SEQ ID NO:26, ADP44073), GABBR1 (578 amino acids, SEQ
CC ID NO:28, ADP44075), P2X1 (150 amino acids, SEQ ID NO:30, ADP44077),
CC HELO1 (299 amino acids, SEQ ID NO:32, ADP44079), GPRK6 (576 amino acids,
CC SEQ ID NO:34, ADP44081), or PTK2B (1009 amino acids, SEQ ID NO:36,
CC ADP44083). Also described: (1) an inhibitor of a cell surface polypeptide
CC as described above in a human host cell preventing HIV entry into the
CC human host cell; (2) a pharmaceutical composition comprising an inhibitor
CC of (1) and a carrier; and (3) conferring resistance to HIV infection in
CC an individual, by administering the pharmaceutical composition of (2). An
CC inhibitor of HIV entry has anti-HIV and virucide activities, and can be
CC used as an HIV uptake inhibitor. (M1) is useful for identifying
CC protective compounds that inhibit entry of HIV into cells, useful for the
CC prevention and/or treatment of HIV infection. The present sequence
CC represents human integrin alpha X (CD11C), which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 1163 AA;

Query Match 58.9%; Score 3460; DB 8; Length 1163;
Best Local Similarity 60.7%; Pred. No. 3.7e-277;
Matches 685; Conservative 143; Mismatches 295; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||    |: :: ||| ||||    | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||||||||||:|||| ||||||||||  |  | |: |||||  |    |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: ||-||||||||||    :|  |    |  :: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              : ||||:||: : || ||:  :  || | |:||| :: ||  ||: : |||::| ||| ::
Db        198 QTHFTFEEFRRSSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAAKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||:| || | |:|||| |  |:||| :| | ||::  | :||| :|||| ::|:|:
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| :|||:|||||||||:| |||||| ||:||||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |  ||||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:||||||||||||:|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDSDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || |  ||||||||||||:|| |||||||:|||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :||| :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || |||:|| |||| ||| ||||||| |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAADAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::|| |||| || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||||| ||:||   ||||||: | |  | |  |:|||:|||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| |  | ::::  |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :     : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 11
ADN02004
ID ADN02004 standard; protein; 1163 AA.
XX
AC ADN02004;
XX
DT 17-JUN-2004 (first entry)
XX
DE Human inflammatory bowel disease marker - CD11c protein.
XX
KW marker gene; inflammatory bowel disease; IBD; FcgammaR IIIa;
KW FcgammaR IIIb; Mig; NRG-2; hexokinase; HM74; REG III; LPAP; Mip-1(beta);
KW L-selectin; EGFL6; IDO; IL-8; CD11c; TLR2; ulcerative colitis;
KW Crohn's disease.
XX
OS Homo sapiens.
XX
PN JP2004065120-A.
XX
PD 04-MAR-2004.
XX
PF 07-AUG-2002; 2002JP-00229705.
XX
PR 07-AUG-2002; 2002JP-00229705.
XX
PA (SUMU) SUMITOMO SEIYAKU KK.
XX
DR WPI; 2004-209124/20.
DR N-PSDB; ADN01990.
XX
PT Novel marker gene of inflammatory bowel disease (IBD) comprising base
PT sequence of e.g., Fc gamma R IIIa, Fc gamma R IIIb, Mig, NRG-2,
PT hexokinase 3, HM74, CD11c, TLR2 gene, useful for screening therapeutic

```
PT   agent for IBD.
XX
PS   Example; SEQ ID NO 27; 151pp; Japanese.
XX
CC   The invention comprises marker genes for inflammatory bowel disease (IBD)
CC   - FcgammaR IIIa, FcgammaR IIIb, Mig, NRG-2, hexokinase, HM74, REG III,
CC   LPAP, Mip-1(beta), L-selectin, EGFL6, IDO, IL-8, CD11c, and TLR2 genes.
CC   The DNA and protein sequences of the invention are useful in the
CC   detection and treatment of IBD (e.g. ulcerative colitis and Crohn's
CC   disease). The present amino acid sequence represents an IBD marker
CC   protein of the invention.
XX
SQ   Sequence 1163 AA;

  Query Match             58.6%;  Score 3446;  DB 8;  Length 1163;
  Best Local Similarity   60.5%;  Pred. No. 5.4e-276;
  Matches  683;  Conservative  143;  Mismatches  297;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||| |: :: ||| |||| | ||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               |||| |||||||||||:|||| |||||||||  |  | |: |||||  |    | |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
               | : : ||::: || |||||||||   :|  |   |  :: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              : ||||:||:   || ||:  : || | |:||| :: ||  ||: : |||::| ||| ::
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||:| || | |:|||| |  |:||| :| | ||::  | :||| :|||| ::|:|:
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| :|||:|||||||||:| |||||| ||:|||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      ||||:: : || |:||||:  : |   ||||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:|||||||||||::|||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || ||||||||||| ||||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  |||||||||||||:|| |||||||:|||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :|||  :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
               ||||| ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:||  ||| |||  ||||||| |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFPFLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |  ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTPNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |  |||| ||| ||  :::|| |||| || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| ||:||  |||||: | |  | |  |:|||:|||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| |  | ::  |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :     : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:|||||||||| |||:||||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143


RESULT 12
ADQ17510
ID   ADQ17510 standard; protein; 1163 AA.
XX
AC   ADQ17510;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated protein - SEQ ID 327.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 327; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   protein of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 1163 AA;
```

Query Match 58.6%; Score 3446; DB 8; Length 1163;
Best Local Similarity 60.5%; Pred. No. 5.4e-276;
Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 13
AAR07120
ID AAR07120 standard; protein; 1163 AA.
XX
AC AAR07120;
XX
DT 25-MAR-2003 (revised)
DT 05-FEB-1991 (first entry)
XX
DE p150.95 alpha subunit encoded by clone lambdaX47.
XX
KW p150,95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;
KW rhinovirus.
XX
OS Synthetic.
XX

FH	Key	Location/Qualifiers
FT	Region	1. .19
FT		/label= signal peptide
FT	Region	20. .44
FT		/label= N-terminus
FT	Modified-site	61. .63
FT		/label= glycosylation site
FT	Modified-site	89. .91
FT		/label= glycosylation site
FT	Modified-site	385. .387
FT		/label= glycosylation site
FT	Modified-site	392. .394
FT		/label= glycosylation site
FT	Modified-site	697. .699
FT		/label= glycosylation site
FT	Modified-site	735. .737
FT		/label= glycosylation site
FT	Modified-site	899. .901
FT		/label= glycosylation site
FT	Modified-site	904. .906
FT		/label= glycosylation site
FT	Modified-site	939. .941
FT		/label= glycosylation site
FT	Modified-site	1050. .1052
FT		/label= glycosylation site
FT	Domain	1108. .1133
FT		/label= transmembrane

XX
PN WO9010646-A.
XX
PD 20-SEP-1990.
XX
PF 09-MAR-1990; 90WO-US001257.
XX
PR 09-MAR-1990; 90WO-US001257.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Corbi AA, Springer TA;
XX
DR WPI; 1990-304985/40.
DR N-PSDB; AAQ06068.
XX
PT Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-
PT unit of p150.95 cell surface adhesion receptor, opt. together with a beta
PT chain of CD-18 family.
XX
PS Disclosure; Fig 3; 59pp; English.
XX

CC Clone lambda X47 was isolated from a cDNA library constructed from total
CC RNA extracted from phorbol myristate acetate stimulated HL-60
CC myelomonocytic cells. The library was screened with oligonucleotide
CC probes based on tryptic peptide fragments of p150.95. The sequence can be
CC attached to appropriate control elements and expressed in prokaryotic and
CC eukaryotic cells. The protein can be used to treat or prevent rhinoviral
CC infection because it interacts with ICAM-1 and inhibits cell-virus
CC attachment. It can also be used as an anti-inflammatory agent. See also
CC AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003
CC to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1163 AA;

Query Match 58.5%; Score 3437; DB 2; Length 1163;
Best Local Similarity 60.3%; Pred. No. 3e-275;
Matches 681; Conservative 143; Mismatches 299; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||    | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               |||| ||||||||||||:|||| |||||||||| |  | |: ||||| |    |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || |||||||||    :|  |   |  :: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              : ||||:||:   || ||:  : || | |:||| :: ||  ||: : |||::| ||| ::
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILILVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||:| || | |:|||  |  |:||| :| | ||::  | :||| :|||| ::|:|:
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| :||||:|||||||||:| |||||| ||:||||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||       |||||:: : || |:||||:  : |   |||||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:||||||||||||::|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || ||||||||| |||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  ||||||||||||:|| |||||||:|||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |    |:| | |: |: ||| :|||  :   :  :||:: | ::: |     :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              |||||  ||    || | |||| :  :  :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:||  ||| |||  ||||||| |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | | | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSA--PVA 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::||  |||| || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| ||:||   ||||||: | |  | |  |:|||:|||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | ||| :|   :|:  :| |  | ::: | |||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSQKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :     : :|| ||| |: ||:: |||  |:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQERFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:| |||||||| |||:||||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGALLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 14
AAW65091
ID AAW65091 standard; protein; 1163 AA.
XX
AC AAW65091;
XX
DT 28-SEP-1998 (first entry)
XX
DE Human Beta-integrin CD11c subunit protein.
XX
KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN US5728533-A.
XX
PD 17-MAR-1998.
XX
PF 07-JUN-1995; 95US-00485618.
XX
PR 23-DEC-1993; 93US-00173497.
PR 05-AUG-1994; 94US-00286889.
PR 21-DEC-1994; 94US-00362652.
XX
PA (ICOS-) ICOS CORP.
XX
PI Van Der Vieren M, Gallatin WM;
XX
DR WPI; 1998-206565/18.
XX
PT Screening assay for modulators of integrin binding - using immobilised or
PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.
XX
PS Example 5; Fig 1A-D; 106pp; English.
XX
CC This sequence represents a human beta-integrin CD11c subunit which is
CC used to describe a method for identifying compounds that modulate the
CC interaction of the beta-integrin alpha-d subunit with a binding partner
CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha
CC -d binding partner, one of which is immobilised and the other of which is
CC labelled, in the presence of a test compound, and determining if the
CC compound affects binding between the alpha-d polypeptide and alpha-d
CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment
CC comprising the cytoplasmic, transmembrane or extracellular domain of
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat
CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC and rheumatoid arthritis
XX

SQ Sequence 1163 AA;

Query Match 58.2%; Score 3423; DB 2; Length 1163;
Best Local Similarity 60.2%; Pred. No. 4.4e-274;
Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||    | |||||||:|:|||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||:|||| ||||||||||  | | |: ||||| |   |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || |||||||||   :|  |   | :: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              : ||||:||:   || ||:  : || | |:||| :: ||  ||: : |||::| ||| ::
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAIKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||:| || | |:|||| |  |:||| :| | ||::  | :||| :|||| ::|:|:
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| :||||:|||||||||:| |||||| ||:||||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||       |||||:: : || |:||||:  : |   |||||||||||||||   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |    | | ||||||:||||||||||||::|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVIGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| ||||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  ||||||||||||:|| |||||||:||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :|||  :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||  ||   ||:| |||| :  : :||||   ||   | ||:|:|| | ||:|||||
Db        677 LDLALAPGRLSPRAIFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || |||||:||  ||| |||  ||||||| |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |  ||  | |  |:
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTC--CSAPVG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::|| ||||  ||   ::||| |||:||||:|||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              ||||||||||:||:||   ||||||: | |  | |  |:|||:||||| ||:|: | ||
Db        915 QLELPVKYAVYIVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| | |  | :::  |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034
```

```
Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : ||||||||| |  :     : :|| |||:|: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEIIFDTSVYSQLPGQEAFMRAQTI 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| ||:|||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPIPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 15
AAB07361
ID AAB07361 standard; protein; 1163 AA.
XX
AC AAB07361;
XX
DT 17-JAN-2001 (first entry)
XX
DE Human CD11c protein sequence.
XX
KW Human; macrophage infiltration inhibition; alpha_d integrin;
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW rheumatoid arthritis; central nervous system injury; CD11c.
XX
OS Homo sapiens.
XX
PN WO200029446-A1.
XX
PD 25-MAY-2000.
XX
PF 16-NOV-1999; 99WO-US027139.
XX
PR 16-NOV-1998; 98US-00193043.
PR 08-JUL-1999; 99US-00350259.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gallatin MW, Van Der Vieren M;
XX
DR WPI; 2000-387751/33.
XX
PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous system
PT injury sites.
XX
PS Example 5; Fig 1; 270pp; English.
XX
CC Integrins are a class of membrane-associated molecules that participate
CC in cellular adhesion. Integrins are made up of an alpha subunit and a
CC beta subunit. One class of human integrins are restricted to expression
CC in white blood cells and have a common beta2 subunit: the leukocyte
CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC have an important role in immune and inflammatory responses. The present
CC protein sequence is the human integrin alpha subunit CD11c. This sequence
CC was used in an alignment to identify a novel beta2 integrin alpha
CC subunit: alpha_d (AAA60014 and AAB07359). The present sequence has
CC approximately 66% identity to the protein sequence of alpha_d. The
CC Alpha_d gene and protein may be useful in therapy for diseases linked to
CC aberrant alpha_d function e.g. Type I diabetes, atherosclerosis, multiple
CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
CC (LAD). In addition, anti-alpha_d monoclonal antibodies may be used in the
CC inhibition of macrophage infiltration at the site of a central nervous
CC system injury. The monoclonal antibodies can also be used to detect and
CC diagnose Crohn's disease
XX
SQ Sequence 1163 AA;

Query Match 58.2%; Score 3423; DB 3; Length 1163;
Best Local Similarity 60.2%; Pred. No. 4.4e-274;

Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAIKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        318 DFDALKDIQNQLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        438 QVSRQWRMKAEVIGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALAPGRLSPRAIFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTC--CSAPVG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYIVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEIIFDTSVYSQLPGQEAFMRAQTI 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPIPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

Search completed: November 9, 2004, 12:08:14
Job time : 109.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 12:17:49 ; Search time 97.5 Seconds
(without alignments)
4119.157 Million cell updates/sec

Title: US-09-902-481B-3
Perfect score: 5879
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5879	100.0	1137	10	US-09-902-481A-3	Sequence 3, Appli
2	5857	99.6	1137	10	US-09-902-481A-4	Sequence 4, Appli
3	5839	99.3	1153	9	US-09-350-259-3	Sequence 3, Appli
4	5839	99.3	1153	10	US-09-902-481A-1	Sequence 1, Appli
5	5839	99.3	1153	10	US-09-891-943-3	Sequence 3, Appli
6	5839	99.3	1153	14	US-10-144-259-30	Sequence 30, Appl
7	5839	99.3	1153	14	US-10-207-655-176	Sequence 176, App
8	5836	99.3	1137	10	US-09-902-481A-5	Sequence 5, Appli
9	5832	99.2	1137	10	US-09-902-481A-6	Sequence 6, Appli
10	5823.5	99.1	1152	9	US-09-945-265-4	Sequence 4, Appli
11	3446	58.6	1163	14	US-10-116-275-204	Sequence 204, App
12	3423	58.2	1163	9	US-09-350-259-4	Sequence 4, Appli
13	3423	58.2	1163	10	US-09-891-943-4	Sequence 4, Appli
14	3388	57.6	1161	9	US-09-350-259-2	Sequence 2, Appli
15	3388	57.6	1161	10	US-09-891-943-2	Sequence 2, Appli
16	3372.5	57.4	1161	9	US-09-350-259-99	Sequence 99, Appl
17	3372.5	57.4	1161	10	US-09-891-943-99	Sequence 99, Appl
18	3213.5	54.7	1161	9	US-09-350-259-55	Sequence 55, Appl
19	3213.5	54.7	1161	10	US-09-891-943-55	Sequence 55, Appl
20	3201	54.4	1161	9	US-09-350-259-53	Sequence 53, Appl
21	3201	54.4	1161	10	US-09-891-943-53	Sequence 53, Appl
22	3192.5	54.3	1151	9	US-09-350-259-37	Sequence 37, Appl
23	3192.5	54.3	1151	10	US-09-891-943-37	Sequence 37, Appl
24	3180	54.1	1155	9	US-09-350-259-46	Sequence 46, Appl
25	3180	54.1	1155	10	US-09-891-943-46	Sequence 46, Appl
26	1848	31.4	369	13	US-10-087-192-1212	Sequence 1212, Ap
27	1536.5	26.1	1170	9	US-09-945-265-2	Sequence 2, Appli
28	1532.5	26.1	1170	15	US-10-261-164-1	Sequence 1, Appli
29	1509	25.7	1223	16	US-10-408-765A-295	Sequence 295, App
30	1341.5	22.8	1086	16	US-10-408-765A-1871	Sequence 1871, Ap
31	1229.5	20.9	494	9	US-09-350-259-103	Sequence 103, App
32	1229.5	20.9	494	10	US-09-891-943-103	Sequence 103, App
33	1128.5	19.2	413	9	US-09-350-259-101	Sequence 101, App
34	1128.5	19.2	413	10	US-09-891-943-101	Sequence 101, App
35	1128	19.2	1179	14	US-10-177-550-2	Sequence 2, Appli
36	1128	19.2	1179	14	US-10-173-551-2	Sequence 2, Appli
37	1087	18.5	1188	14	US-10-291-265-810	Sequence 810, App
38	1085.5	18.5	1151	10	US-09-984-130-103	Sequence 103, App
39	1085.5	18.5	1151	10	US-09-836-353A-103	Sequence 103, App
40	1085.5	18.5	1179	10	US-09-918-715-250	Sequence 250, App
41	1085.5	18.5	1179	17	US-10-474-794-250	Sequence 250, App
42	1083	18.4	1188	14	US-10-291-265-338	Sequence 338, App
43	1079.5	18.4	1189	10	US-09-984-130-35	Sequence 35, Appl
44	1079.5	18.4	1189	10	US-09-836-353A-35	Sequence 35, Appl
45	1079.5	18.4	1189	15	US-10-262-839-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 100.0%; Score 5879; DB 10; Length 1137;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
```

```
Qy       121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
Db       121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180

Qy       181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db       181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Qy       241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db       241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300

Qy       301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db       301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy       361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db       361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy       421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db       421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy       481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db       481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy       541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db       541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy       601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db       601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy       661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db       661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy       721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db       721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy       781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db       781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy       841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db       841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy       901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db       901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy       961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db       961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy      1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db      1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy      1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db      1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

RESULT 2
US-09-902-481A-4
; Sequence 4, Application US/09902481A

; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4

Query Match 99.6%; Score 5857; DB 10; Length 1137;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1131; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy         1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy        61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db        61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Qy       121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
Db       121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180

Qy       181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db       181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Qy       241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db       241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Qy       301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db       301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy       361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db       361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy       421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db       421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy       481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db       481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy       541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db       541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy       601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db       601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy       661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db       661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
```

```
Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

RESULT 3
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.3%; Score 5839; DB 9; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||||||||| ||||:||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              ||||||||||||||||| |||||||||:| ||||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 4
US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy

```
;  APPLICANT:  Shimaoka, Motomu
;  APPLICANT:  Shifman, Julia
;  APPLICANT:  Mayo, Stephen
;  TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
;  FILE REFERENCE: A-70586-1/RFT/RMS/RMK
;  CURRENT APPLICATION NUMBER: US/09/902,481A
;  CURRENT FILING DATE:  2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/216,600
;  PRIOR FILING DATE: 2000-07-07
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (17)..()
;   OTHER INFORMATION:
US-09-902-481A-1

  Query Match             99.3%;  Score 5839;  DB 10;  Length 1153;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||||||||||||| |||:||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||:||||||||||||||:|||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||| |||||||||:| ||||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 5
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT:  Van der Vieren, Monica
;  TITLE OF INVENTION: No. US20030077278A1el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/891,943
;  CURRENT FILING DATE:  2001-06-26
;  PRIOR APPLICATION NUMBER: 09/193,043
;  PRIOR FILING DATE: 1998-11-16
;  PRIOR APPLICATION NUMBER: 08/286,889
;  PRIOR FILING DATE: 1994-08-05
;  PRIOR APPLICATION NUMBER: 08/362,652
;  PRIOR FILING DATE: 1994-12-21
;  PRIOR APPLICATION NUMBER: 08/943,363
;  PRIOR FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-891-943-3

  Query Match             99.3%;  Score 5839;  DB 10;  Length 1153;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||||||||||||| |||:||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
```

```
                  |||||||||||||||||:|||||||||||||||:||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
                  |||||||||||||||| |||||||||:| |||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
                  |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 6
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
;  APPLICANT: Arnaout, M. Amin
;  APPLICANT:  Li, Rui
;  APPLICANT:  Xiong, Jian-Ping
;  TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: 00786-548001
;  CURRENT APPLICATION NUMBER: US/10/144,259
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: US 09/758,493
;  PRIOR FILING DATE: 2001-01-11
;  PRIOR APPLICATION NUMBER: US 60/221,950
;  PRIOR FILING DATE: 2000-07-31
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-144-259-30

  Query Match             99.3%;  Score 5839;  DB 14;  Length 1153;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||| |||:||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||| |||||||||:| |||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 7
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 99.3%; Score 5839; DB 14; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180

Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300

Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 8
US-09-902-481A-5
; Sequence 5, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic

US-09-902-481A-5

Query Match 99.3%; Score 5836; DB 10; Length 1137;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||| || :||:||||||||||||||||||||
Db        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||| |:
Db        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||| |||||||||:| ||||||||||||||||||||||||||| |
Db        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

RESULT 9
US-09-902-481A-6
; Sequence 6, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6

Query Match 99.2%; Score 5832; DB 10; Length 1137;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1122; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||||| |||:||||||||||||||||||
Db        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:|||||||||||||||:|||:||| |||||||||||||| ::
Db        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||| |||||||||:| ||||||||||||||||||||||||||| |
Db        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137


RESULT 10
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
;  APPLICANT: Springer, Timothy A.
;  APPLICANT:  Shimaoka, Motomu
;  APPLICANT:  Lu, Chafen
;  TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
;  TITLE OF INVENTION:  DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
;  FILE REFERENCE: CBN-002CP
;  CURRENT APPLICATION NUMBER: US/09/945,265
;  CURRENT FILING DATE:  2001-08-31
;  PRIOR APPLICATION NUMBER: US 60/229,700
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 1152
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-945-265-4

  Query Match            99.1%;  Score 5823.5;  DB 9;  Length 1152;
  Best Local Similarity  98.8%;  Pred. No. 0;
  Matches 1123;  Conservative    8;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180

Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300

Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSPEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152
```

RESULT 11
US-10-116-275-204
; Sequence 204, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 204
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-204

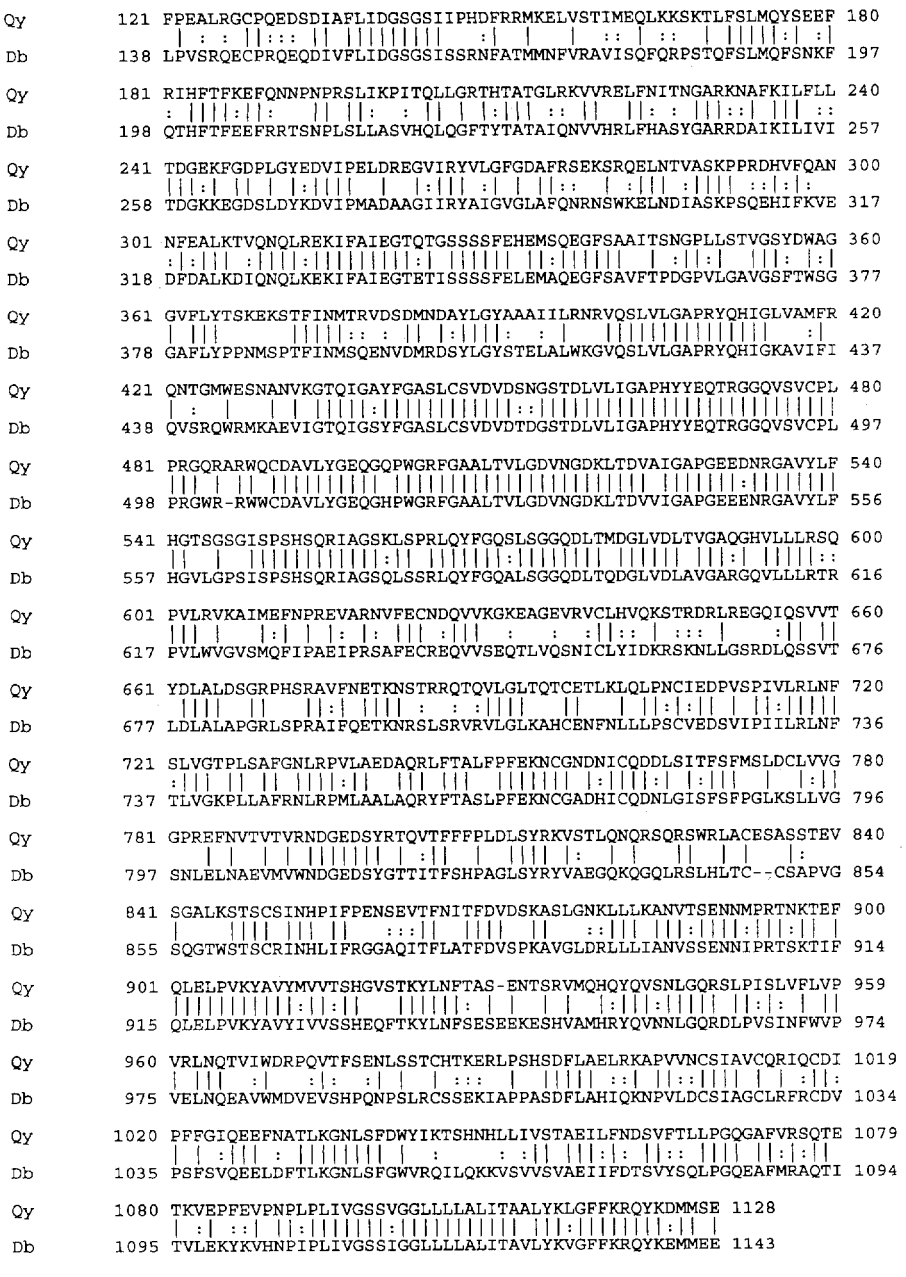
Query Match 58.6%; Score 3446; DB 14; Length 1163;
Best Local Similarity 60.5%; Pred. No. 1.2e-284;
Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||   | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||||||||||:|||| ||||||||||   |  | |: ||||| |     |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || |||||||||   :|  |   |  :: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              : ||||:||:   || ||:  : || | |:||| :: ||  ||: : |||::| ||| ::
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||:| || | |:|||| |  |:||| :| | ||::  | :||| :|||| ::|:|:
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| :||||:|||||||||:| |||||| ||:||||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||       ||||:: : || |:||||:  : |   |||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:||||||||||||::||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  ||||||||||:|| |||||||:|||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :||| :   :  :||:: | ::: |     :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:||  ||| |||  ||||||| |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | | | | |||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||  :::||  ||||  || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||||| ||:||  ||||||: | |  | |  |:|||:|||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| | | ::  |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :     : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 12
US-09-350-259-4
; Sequence 4, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-4

Query Match 58.2%; Score 3423; DB 9; Length 1163;
Best Local Similarity 60.2%; Pred. No. 1.1e-282;
Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;

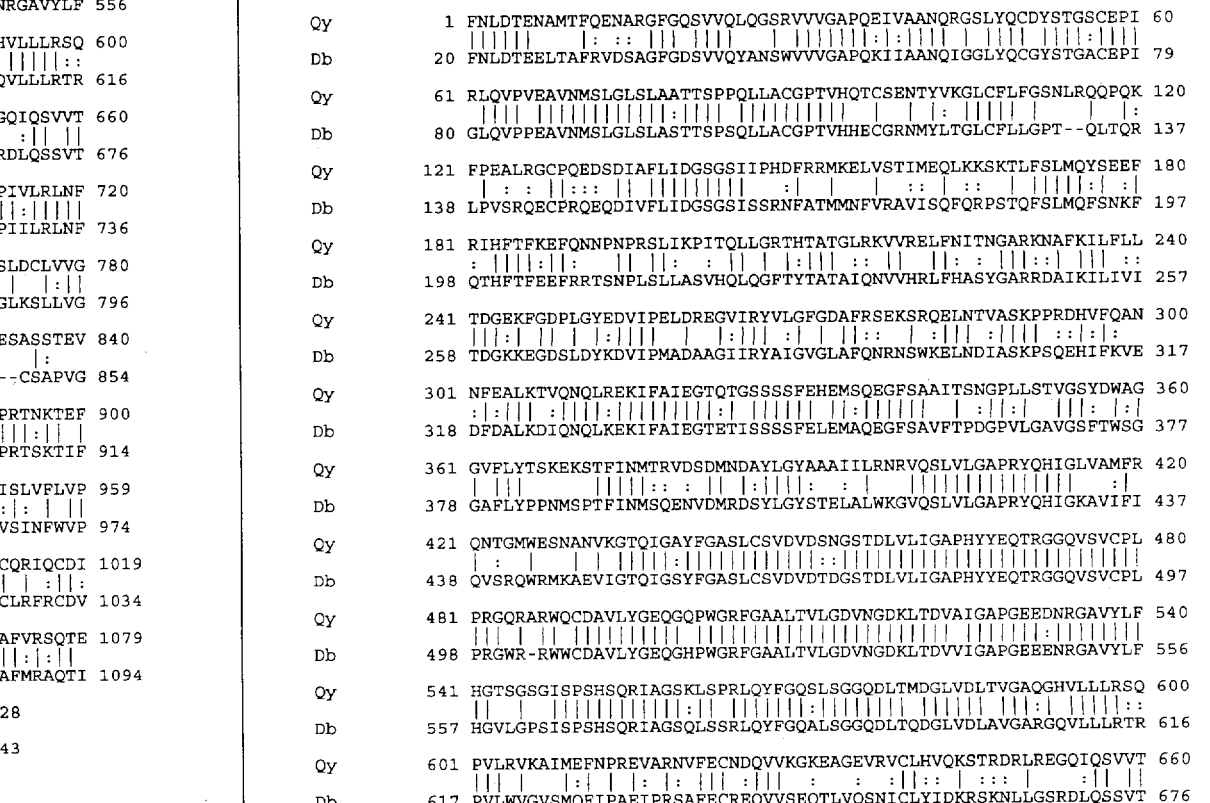
```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||   | |||||||:|:|||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||||||||||:|||| ||||||||||   |  | |: ||||| |     |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137
```



RESULT 13
US-09-891-943-4
; Sequence 4, Application US/09891943

; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-4

Query Match 58.2%; Score 3423; DB 10; Length 1163;
Best Local Similarity 60.2%; Pred. No. 1.1e-282;
Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;



```
Db        677 LDLALAPGRLSPRAIFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTC--CSAPVG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959

Db        915 QLELPVKYAVYIVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019

Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079

Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEIIFDTSVYSQLPGQEAFMRAQTI 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128

Db       1095 TVLEKYKVHNPIPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 14
US-09-350-259-2
; Sequence 2, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-2

Query Match 57.6%; Score 3388; DB 9; Length 1161;
Best Local Similarity 58.9%; Pred. No. 1.1e-279;
Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQTGRLYDCAAATGMCQPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-EIIQT 135

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180

Db        136 VPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFEGTDTLFALMQYSNLL 195

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Db        196 KIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHHKNGARKSAKKILIVI 255

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300

Db        256 TDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELNTISSAPPQDHVFKVD 315

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        316 NFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGSFSWSG 375

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        376 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVLGAPRYQHTGKAVIFT 435

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        436 QVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPHYYEQTRGGQVSVCPL 495

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        496 PRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAIGAPGEQENRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLLRSL 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        616 PVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        674 FDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCVEDVVSPIILHLNF 733

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        734 SLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDLGVTLSFSGLQTLTVG 793

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        794 SSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        853 EG-LRSSRCSVNHPIFHEGSNGTFIVTFDVSYKATLGDRMLMRASASSENNKASSSKATF 911

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNF-TASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959

Db        912 QLELPVKYAVYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQRDLAISINFWVP 971

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019

Db        972 VLLNGVAVWDVVMEAPSQSL--PCVSERKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1029

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079

Db       1030 PSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQLPGQEAFMRAQME 1089

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128

Db       1090 MVLEEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRHYKEMLED 1138
```

RESULT 15
US-09-891-943-2
; Sequence 2, Application US/09891943
; Publication No. US20030077278A1

; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-2

Query Match 57.6%; Score 3388; DB 10; Length 1161;
Best Local Similarity 58.9%; Pred. No. 1.1e-279;
Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    |||:| |||||||  |||:|||| |:|||| | || |  :|| |:||
Db         17 FNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQTGRLYDCAAATGMCQPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               | :  |||||||||:|||:|:  :|||||||:|: | ||:| || | | ||   : |
Db         77 PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-EIIQT 135

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |:|   || :: || |||||||||  :|| :||  |  :| | : : |||:|||||
Db        136 VPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFEGTDTLFALMQYSNLL 195

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              :||||| :|: :|: :||: || || | | ||||:  || :||:  |||||:| ||| ::
Db        196 KIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHHKNGARKSAKKILIVI 255

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||:|: ||| | ||||: :: |:||| :| | ||:   :||||||::| ||:||||: :
Db        256 TDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELNTISSAPPQDHVFKVD 315

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              || || ::| ||:|||:|:|||| :||||:|||||||| |:| :|  |  |||: |:|
Db        316 NFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGSFSWSG 375

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |   ||:|||||||||| |   :|
Db        376 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVLGAPRYQHTGKAVIFT 435

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:||||||||||||:|||||:|||||||||||||||||||||
Db        436 QVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPHYYEQTRGGQVSVCPL 495

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||| :|||||| |||| ||||||||||||||| ||| |||||||||::||||||||
Db        496 PRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAIGAPGEQENRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || | |||||||||||| }:|||||||||:|||||||| |||:|| |||:| ||||||
Db        556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLLRSL 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |||:|   | |:| |||: |: | ::     |||:  ||| :|||: |:|  | ||| |
Db        616 PVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              :||||| ||  |||:|||||| |  : : |||   |||||| ||:|:|| |||||:| |||

Db        674 FDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCVEDVVSPIILHLNF 733

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||  |: :  |||||||  :| |||| |||||||  | :|: || :| || |  | ||
Db        734 SLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDLGVTLSFSGLQTLTVG 793

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | || ||| | ||||| | |: ::|  ||:|:|| | |  | : |||||:   ||
Db        794 SSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              | |:|: ||:||||| | |  || :|||| ||:||:::|::|: :||||   ::|  |
Db        853 EG-LRSSRCSVNHPIFHEGSNGTFIVTFDVSYKATLGDRMLMRASASSENNKASSSKATF 911

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNF-TASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| :::    |||| || |: |   :  :|:|:|:|| || | ||: | ||
Db        912 QLELPVKYAVYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQRDLAISINFWVP 971

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | ||   :||       |::|   | :: : | ||||| :: ::|:::|||| | : :||:
Db        972 VLLNGVAVWDVVMEAPSQSL--PCVSERKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1029

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :|    :|:|| ||| |: ||:: |||| ||:|:| |
Db       1030 PSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQLPGQEAFMRAQME 1089

Qy       1080 TKVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
                :|  || | :|:|:||||| |||||||| |||||||| ||:|: :
Db       1090 MVLEEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRHYKEMLED 1138
```

Search completed: November 9, 2004, 12:46:48
Job time : 102.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:57 ; Search time 23 Seconds
(without alignments)
3278.416 Million cell updates/sec

Title: US-09-902-481B-3
Perfect score: 5879
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5839	99.3	1153	1	US-08-173-497-3	Sequence 3, Appli
2	5839	99.3	1153	1	US-08-286-889-3	Sequence 3, Appli
3	5839	99.3	1153	1	US-08-485-618-3	Sequence 3, Appli
4	5839	99.3	1153	1	US-08-362-652-3	Sequence 3, Appli
5	5839	99.3	1153	2	US-08-605-672-3	Sequence 3, Appli
6	5839	99.3	1153	2	US-08-482-293A-3	Sequence 3, Appli
7	5839	99.3	1153	2	US-08-943-363-3	Sequence 3, Appli
8	5839	99.3	1153	3	US-09-193-043-3	Sequence 3, Appli
9	5839	99.3	1153	4	US-09-688-307A-3	Sequence 3, Appli
10	5839	99.3	1153	4	US-09-350-259-3	Sequence 3, Appli
11	5808.5	98.8	1152	2	US-08-476-062A-43	Sequence 43, Appl
12	5808.5	98.8	1152	5	PCT-US96-01314-43	Sequence 43, Appl
13	5808.5	98.8	1152	6	5424399-2	Patent No. 5424399
14	3446	58.6	1163	2	US-08-476-062A-44	Sequence 44, Appl
15	3446	58.6	1163	5	PCT-US96-01314-44	Sequence 44, Appl
16	3423	58.2	1163	1	US-08-173-497-4	Sequence 4, Appli
17	3423	58.2	1163	1	US-08-286-889-4	Sequence 4, Appli
18	3423	58.2	1163	1	US-08-485-618-4	Sequence 4, Appli
19	3423	58.2	1163	1	US-08-362-652-4	Sequence 4, Appli
20	3423	58.2	1163	2	US-08-605-672-4	Sequence 4, Appli
21	3423	58.2	1163	2	US-08-482-293A-4	Sequence 4, Appli
22	3423	58.2	1163	2	US-08-943-363-4	Sequence 4, Appli
23	3423	58.2	1163	3	US-09-193-043-4	Sequence 4, Appli
24	3423	58.2	1163	4	US-09-688-307A-4	Sequence 4, Appli
25	3423	58.2	1163	4	US-09-350-259-4	Sequence 4, Appli
26	3388	57.6	1161	1	US-08-173-497-2	Sequence 2, Appli
27	3388	57.6	1161	1	US-08-286-889-2	Sequence 2, Appli
28	3388	57.6	1161	1	US-08-485-618-2	Sequence 2, Appli
29	3388	57.6	1161	1	US-08-362-652-2	Sequence 2, Appli
30	3388	57.6	1161	2	US-08-605-672-2	Sequence 2, Appli
31	3388	57.6	1161	2	US-08-482-293A-2	Sequence 2, Appli
32	3388	57.6	1161	2	US-08-943-363-2	Sequence 2, Appli
33	3388	57.6	1161	3	US-09-193-043-2	Sequence 2, Appli
34	3388	57.6	1161	4	US-09-688-307A-2	Sequence 2, Appli
35	3388	57.6	1161	4	US-09-350-259-2	Sequence 2, Appli
36	3372.5	57.4	1161	1	US-08-485-618-99	Sequence 99, Appl
37	3372.5	57.4	1161	2	US-08-605-672-99	Sequence 99, Appl
38	3372.5	57.4	1161	2	US-08-482-293A-99	Sequence 99, Appl
39	3372.5	57.4	1161	2	US-08-943-363-99	Sequence 99, Appl
40	3372.5	57.4	1161	3	US-09-193-043-99	Sequence 99, Appl
41	3372.5	57.4	1161	4	US-09-688-307A-99	Sequence 99, Appl
42	3372.5	57.4	1161	4	US-09-350-259-99	Sequence 99, Appl
43	3213.5	54.7	1161	3	US-09-193-043-55	Sequence 55, Appl
44	3213.5	54.7	1161	4	US-09-688-307A-55	Sequence 55, Appl
45	3213.5	54.7	1161	4	US-09-350-259-55	Sequence 55, Appl

ALIGNMENTS

```
RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van Der Vieren, Monica
;    TITLE OF INVENTION:  No. 5437958el Human  2 Integrin Alpha
;    TITLE OF INVENTION:  Subunit
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 S. Wacker Drive, 6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/173,497
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5437958and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/31363
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-173-497-3

  Query Match             99.3%;  Score 5839;  DB 1;  Length 1153;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76
```

```
Qy        61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy       121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
             |||||||||||||||||||||||||||||||||||||||||||| |||:||||||||||||||
Db       137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy       181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
             |||||||||||||||||||:|||||||||||||||||:|||||||||||||||||||||| ::
Db       197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy       241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
             |||||||||||||||| ||||||||:| ||||||||||||||:|||||||||||| |
Db       257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy       301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
             |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy       361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy       421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy       481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy       541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy       601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy       661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy       721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy       781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy       841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy       901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy       961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy      1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy      1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Mich
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5470953el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,889
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  P38,659
;      REFERENCE/DOCKET NUMBER:  27866/32168
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-286-889-3

  Query Match             99.3%;  Score 5839;  DB 1;  Length 1153;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||||||||||||| |||:||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||||:|||||||||||||||||:|||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||| ||||||||:| ||||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376
```

```
Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 3
US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5728533el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,618
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32797
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-485-618-3

  Query Match             99.3%;  Score 5839;  DB 1;  Length 1153;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||| |||:|||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||| |||||||:| |||||||||||||:|||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
```

```
                                                
```

```
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 5
US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5817515el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/605,672
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32684
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-605-672-3
```

```
Query Match           99.3%;  Score 5839;  DB 2;  Length 1153;
Best Local Similarity  98.9%;  Pred. No. 0;
Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180

Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300

Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036
```

```
Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 6
US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5831029el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,293A
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32684
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-482-293A-3

  Query Match             99.3%;  Score 5839;  DB 2;  Length 1153;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||||:|||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||| |||||||||: ||||||||||||||:|||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5837478el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/943,363
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32684
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-943-363-3

  Query Match             99.3%;  Score 5839;  DB 2;  Length 1153;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||| |||:||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:||||||||||||||:||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||| ||||||||||:| |||||||||||||||:|||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT: Van der Vieren, Monica
;  TITLE OF INVENTION: No. 6251395el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/193,043
;  CURRENT FILING DATE: 1998-11-16
;  EARLIER APPLICATION NUMBER: 08/173,497
;  EARLIER FILING DATE: 1993-12-23
;  EARLIER APPLICATION NUMBER: 08/286,889
;  EARLIER FILING DATE: 1994-08-05
;  EARLIER APPLICATION NUMBER: 08/362,652
;  EARLIER FILING DATE: 1994-12-21
;  EARLIER APPLICATION NUMBER: 08/943,363
;  EARLIER FILING DATE: 1997-10-03
```

```
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-193-043-3

  Query Match             99.3%;  Score 5839;  DB 3;  Length 1153;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||| |||:||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||| |||||||||:| ||||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 9
US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT:  Van der Vieren, Monica
;  TITLE OF INVENTION: No. 6432404el Human Beta-2
;  FILE REFERENCE: 27866/36646
;  CURRENT APPLICATION NUMBER: US/09/688,307A
;  CURRENT FILING DATE:  2000-10-13
;  PRIOR APPLICATION NUMBER: 09/193,043
;  PRIOR FILING DATE: 1998-11-16
;  PRIOR APPLICATION NUMBER: 08/605,672
;  PRIOR FILING DATE: 1996-02-22
;  PRIOR APPLICATION NUMBER: 08/173,497
;  PRIOR FILING DATE: 1993-12-23
;  PRIOR APPLICATION NUMBER: 08/286,889
;  PRIOR FILING DATE: 1994-08-05
;  PRIOR APPLICATION NUMBER: 08/362,652
;  PRIOR FILING DATE: 1994-12-21
;  PRIOR APPLICATION NUMBER: 08/943,363
;  PRIOR FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-688-307A-3

  Query Match             99.3%;  Score 5839;  DB 4;  Length 1153;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||| |||:||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||| |||||||||:| ||||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 10
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT: Van der Vieren, Monica
;  TITLE OF INVENTION: No. 6620915el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/350,259
;  CURRENT FILING DATE: 1999-07-08
;  EARLIER APPLICATION NUMBER: 09/193,043
;  EARLIER FILING DATE: 1998-11-16
;  EARLIER APPLICATION NUMBER: 08/173,497
;  EARLIER FILING DATE: 1993-12-23
;  EARLIER APPLICATION NUMBER: 08/286,889
;  EARLIER FILING DATE: 1994-08-05
;  EARLIER APPLICATION NUMBER: 08/362,652
;  EARLIER FILING DATE: 1994-12-21
;  EARLIER APPLICATION NUMBER: 08/943,363
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-350-259-3

  Query Match            99.3%;  Score 5839;  DB 4;  Length 1153;
  Best Local Similarity  98.9%;  Pred. No. 0;
  Matches 1124;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||:|||||||||||||||:|||||||||||||||:|||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 11
US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
;  GENERAL INFORMATION:
;    APPLICANT:  Arnaout, M. Amin
;    TITLE OF INVENTION:  CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
;    TITLE OF INVENTION:  RESPONSES WITH BETA2 INTEGRINS
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,062A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/216,081
;      FILING DATE:  21-MAR-1994
;      APPLICATION NUMBER:  07/637,830
;      FILING DATE:  04-JAN-1991
;      APPLICATION NUMBER:  07/539,842
;      FILING DATE:  18-JUN-1990
;      APPLICATION NUMBER:  07/212,573
;      FILING DATE:  28-JUN-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Freeman, John W.
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  00786/068003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1152 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-476-062A-43

  Query Match             98.8%;  Score 5808.5;  DB 2;  Length 1152;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 1121;  Conservative    8;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180

Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300

Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        796 GPRESNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

```
                                                
```

```
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 1121;  Conservative    8;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||| |||:|||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:|||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
              |||||||||||||||||| |||||||||:| |||||||||||||:|||||||||||| |
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPRESNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 14
US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
;  GENERAL INFORMATION:
;    APPLICANT:  Arnaout, M. Amin
;    TITLE OF INVENTION:  CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
;    TITLE OF INVENTION:  RESPONSES WITH BETA2 INTEGRINS
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,062A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/216,081
;      FILING DATE:  21-MAR-1994
;      APPLICATION NUMBER:  07/637,830
;      FILING DATE:  04-JAN-1991
;      APPLICATION NUMBER:  07/539,842
;      FILING DATE:  18-JUN-1990
;      APPLICATION NUMBER:  07/212,573
;      FILING DATE:  28-JUN-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Freeman, John W.
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  00786/068003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1163 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-476-062A-44

  Query Match             58.6%;  Score 3446;  DB 2;  Length 1163;
  Best Local Similarity   60.5%;  Pred. No. 5.1e-278;
  Matches  683;  Conservative  143;  Mismatches  297;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||   | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||||||||||:|||| ||||||||||   |  | |: ||||| |   | |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || |||||||||   :|  |   |  :: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
```

```
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300

Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959

Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019

Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079

Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128

Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

```
RESULT 15
PCT-US96-01314-44
; Sequence 44, Application PC/TUS9601314
;  GENERAL INFORMATION:
;     APPLICANT:  M. Amin Arnaout
;     TITLE OF INVENTION:  METHODS FOR IDENTIFYING INTEGRIN
;     TITLE OF INVENTION:  ANTAGONISTS
;     NUMBER OF SEQUENCES:  78
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fish & Richardson P.C.
;       STREET:  225 Franklin Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  U.S.A.
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;       COMPUTER:  IBM PS/2 Model 50Z or 55SX
;       OPERATING SYSTEM:  MS-DOS (Version 5.0)
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US96/01314
;       FILING DATE:  30-JAN-96
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/380,167
;       FILING DATE:  30-JAN-95
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  John W. Freeman
;       REGISTRATION NUMBER:  29,066
;       REFERENCE/DOCKET NUMBER:  00786/267001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 542-5070
;       TELEFAX:  (617) 542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  44:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1163
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
PCT-US96-01314-44

  Query Match             58.6%;  Score 3446;  DB 5;  Length 1163;
  Best Local Similarity   60.5%;  Pred. No. 5.1e-278;
  Matches  683;  Conservative  143;  Mismatches  297;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180

Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300

Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556
```

```
Qy       541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db       557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy       601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db       617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy       661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db       677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy       721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db       737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy       781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db       797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy       841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db       855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy       901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db       915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy       960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db       975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy      1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db      1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy      1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db      1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

Search completed: November 9, 2004, 12:20:59
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:37 ; Search time 27.25 Seconds
(without alignments)
4014.622 Million cell updates/sec

Title: US-09-902-481B-4
Perfect score: 5884
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5852	99.5	1153	1	RWHU1B	cell surface glyco
2	4460	75.8	1153	2	S00551	leukocyte surface
3	3469	59.0	1163	1	RWHU1C	cell surface glyco
4	1543.5	26.2	1170	2	S03308	cell surface glyco
5	1527.5	26.0	1163	2	I56126	lymphocyte fuction
6	1140	19.4	1179	2	A53213	integrin alpha-E c
7	1093.5	18.6	1151	2	A45226	integrin alpha-1 c
8	1071	18.2	1170	2	I45914	integrin alpha 2 s
9	1057	18.0	1178	2	S44142	VLA-2 protein homo
10	1055	17.9	1180	2	A35854	integrin alpha-1 c
11	1054	17.9	1181	2	A33998	integrin alpha-2 c
12	663	11.3	1039	2	A41131	lymphocyte-Peyer's
13	642	10.9	1038	2	S06046	integrin alpha-4 c
14	630	10.7	1035	2	I58409	integrin alpha-9 c
15	614.5	10.4	1041	2	T31437	integrin alpha cha
16	579.5	9.8	1054	2	JC7294	alphaP integrin -
17	572.5	9.7	1051	2	A35761	cell surface glyco
18	567.5	9.6	1053	2	I55534	VLA-3 alpha subuni
19	555.5	9.4	1053	2	S44250	integrin alpha-5 c
20	546.5	9.3	1034	2	A36108	integrin alpha-V c
21	535	9.1	1044	2	T10050	integrin alpha-v c
22	534.5	9.1	1072	2	A38457	integrin alpha-6 c
23	532	9.0	1049	2	A27079	fibronectin recept
24	532	9.0	1073	2	B36429	integrin alpha-6 c
25	530.5	9.0	1051	2	A40021	integrin alpha-3 c
26	526	8.9	1048	2	A27421	integrin alpha-5 c
27	525.5	8.9	1091	2	A41543	integrin alpha-6 c
28	517	8.8	1044	2	S16516	integrin alpha-8 c
29	496.5	8.4	1146	2	S40311	integrin - fruit f
30	495.5	8.4	1039	2	A34269	integrin alpha-2b
31	494	8.4	1394	2	A29637	position-specific
32	489	8.3	1137	2	JC5950	integrin alpha-7 c
33	488	8.3	1135	2	I61186	alpha-7 integrin -
34	486.5	8.3	1037	2	A60163	glycoprotein IIb -
35	486	8.3	126	2	B30892	leukocyte adhesion
36	470	8.0	1106	2	S38783	integrin alpha cha
37	462	7.9	1226	2	S44824	F54F2.1 protein -
38	453	7.7	1045	2	S60571	integrin alpha v c
39	444.5	7.6	1139	2	S28277	hypothetical prote
40	426	7.2	1115	2	T09403	integrin alpha cha
41	421.5	7.2	1115	2	T09433	integrin alpha cha
42	391	6.6	764	2	I36916	glycoprotein IIb -
43	309.5	5.3	1086	2	T18523	integrin alpha cha
44	299	5.1	604	2	I36917	glycoprotein IIb -
45	286.5	4.9	272	2	A55348	integrin alpha-1 -

ALIGNMENTS

RESULT 1
RWHU1B
cell surface glycoprotein CD11b precursor [validated] - human
N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein M
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567
R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.
J. Biol. Chem. 263, 12403-12411, 1988
A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, C
B.
A;Reference number: A31108; MUID:88315033; PMID:2457584
A;Accession: A31108
A;Molecule type: mRNA
A;Residues: 1-1153 <COR>
A;Cross-references: UNIPROT:P11215; GB:J03925; NID:g187284; PIDN:AAA59544.1; PID:g30714
A;Note: part of this sequence was confirmed by protein sequencing
R;Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
J. Cell Biol. 106, 2153-2158, 1988
A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor
A;Reference number: A28915; MUID:88257215; PMID:2454931
A;Accession: A28915
A;Molecule type: mRNA
A;Residues: 1-499,501-965,'P',967-1153 <ARN>
A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AAA59
A;Note: the authors translated the codon TAC for residue 1129 as Thr
A;Note: part of this sequence, including the amino end of the mature protein, was confi
R;Shelley, C.S.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991
A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally re
A;Reference number: A41600; MUID:92073318; PMID:1683702
A;Accession: A41600
A;Molecule type: DNA
A;Residues: 1-9 <SHE>
A;Cross-references: GB:M76724; NID:g180018; PIDN:AAA58410.1; PID:g553215
R;Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhes
A;Reference number: A94193; MUID:88190151; PMID:2833753
A;Accession: A30892
A;Molecule type: mRNA
A;Residues: 917-1042 <AR2>
A;Cross-references: GB:M18044
R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
A;Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recepto
A;Reference number: A32218; MUID:89098893; PMID:2563162
A;Accession: A32218
A;Molecule type: mRNA
A;Residues: 9-1153 <HIC>
A;Cross-references: GB:J04145; NID:g189068; PIDN:AAA59903.1; PID:g386975

```
A;Note: part of this sequence was confirmed by protein sequencing
R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A;Reference number: A46526; MUID:93123748; PMID:8419480
A;Accession: A46526
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-499,501-1153 <FLE>
A;Cross-references: GB:S52227; NID:g263047; PIDN:AAB24821.1; PID:g263049
A;Note: the last three bases of intron 13, CAG, are included in some but not all mature
A;Note: sequence extracted from NCBI backbone (NCBIP:121963)
R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A;Title: N-terminal sequence of human leukocyte glycoprotein Mo1: conservation across sp
A;Reference number: A90664; MUID:87076671; PMID:3539202
A;Accession: A26091
A;Molecule type: protein
A;Residues: 17-31 <PIE>
A;Experimental source: granulocytes
R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A;Title: Characterization of the myeloid-specific CD11b promoter.
A;Reference number: I52567; MUID:92144986; PMID:1346576
A;Accession: I52567
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-9 <RES>
A;Cross-references: GB:M84477; NID:g180184; PIDN:AAA51960.1; PID:g553219
C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C;Genetics:
A;Gene: GDB:ITGAM; CR3A
A;Cross-references: GDB:120599; OMIM:120980
A;Map position: 16p11.2-16p11.2
A;Note: promoter contains a GATA motif and two Sp1 consensus binding sites
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F;17-1108/Domain: extracellular #status predicted <EXT>
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;465-473/Region: calcium/magnesium binding #status predicted
F;530-538/Region: calcium/magnesium binding #status predicted
F;593-601/Region: calcium/magnesium binding #status predicted
F;1109-1134/Domain: transmembrane #status predicted <TMM>
F;1135-1153/Domain: intracellular #status predicted <INT>
F;86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

  Query Match           99.5%;  Score 5852;  DB 1;  Length 1153;
  Best Local Similarity  99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:||||||||||||:||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 2
S00551
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
N;Alternate names: complement-3 receptor alpha chain
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S00551; I59078
R;Pytela, R.
EMBO J. 7, 1371-1378, 1988
A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the
A;Reference number: S00551; MUID:88312584; PMID:3044779
A;Accession: S00551
A;Molecule type: DNA
A;Residues: 1-1153 <PYT>
A;Cross-references: UNIPROT:P05555; EMBL:X07640; NID:g52982; PIDN:CAA30479.1; PID:g529
A;Note: the authors translated the codon CAC for residue 569 as Gln
```

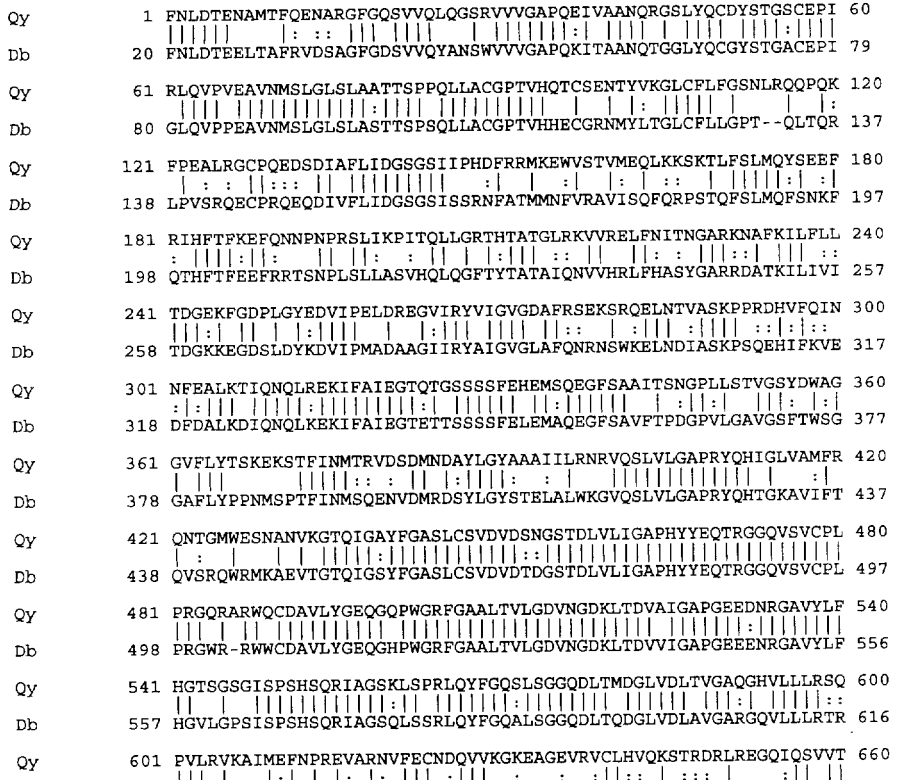
R;Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts,
Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A;Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
A;Reference number: I59078; MUID:86287312; PMID:2942940
A;Accession: I59078
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 11-44 <RES>
A;Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
C;Genetics:
A;Gene: Mac-1
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C;Keywords: cell adhesion; glycoprotein; transmembrane protein
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1106-1129/Domain: transmembrane #status predicted <TMM>

Query Match 75.8%; Score 4460; DB 2; Length 1153;
Best Local Similarity 73.8%; Pred. No. 2.4e-302;
Matches 840; Conservative 146; Mismatches 150; Indels 2; Gaps 2;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         17 FNLDTEHPMTFQENAKGFGQNVVQLGGTSVVVAAPQEAKAVNQTGALYQCDYSTSRCHPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         77 PLQVPPEAVNMSLGLSLAVSTVPQQLLACGPTVHQNCKENTYVNGLCYLFGSNLLRPPQQ 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        137 FPEALRECPQQESDIVFLIDGSGSINNIDFQKMKEFVSTVMEQFKKSKTLFSLMQYSDEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db        197 RIHFTFNDFKRNPSPRSHVSPIKQLNGRTKTASGIRKVVRELFHKTNGARENAAKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db        257 TDGEKFGDPLDYKDVIPEADRAGVIRYVIGVGNAFNKPQSRRELDTIASKPAGEHVFQVD 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        317 NFEALNTIQNQLQEKIFAIEGTQTGSTSSFEHEMSQEGFSASITSNGPLLGSVGSFDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        377 GAFLYTSKDKVTFINTTRVDSDMNDAYLGYASAVILRNRVQSLVLGAPRYQHIGLVVMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        437 ENFGTWEPHTSIKGSQIGSYFGASLCSVDMDADGNTNLILIGAPHYYEKTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        497 PRG-RARWQCEALLHGDQGHPWGRFGAALTVLGDVNGDKLTDVAIGAPGEQENQGAVYIF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        556 YGASIASLSASHSHRIIGAHFSPGLQYFGQSLSGGKDLTMDGLMDLAVGAQGHLLLLRAQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        616 PVLRLEATMEFSPKKVARSVFACQEQVLKNKDAGEVRVCLRVRKNTKDRLREGDIQSTVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        676 YDLALDPVRSRIRAFFDETKNNTRRRTQVFGLMQKCETLKLILPDCVDDSVSPIILRLNY 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        736 TLVGEPLRSFGNLRPVLAMDAQRFFTAMFPFEKNCGNDSICQDDLSITMSAMGLDTLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRL-ACESASSTE 839
Db        796 GPQDFNMSVTLRNDGEDSYGTQVTVYYPSGLSYRKDSASQNPLTKKPWFVKPAESSSSSE 855

Qy        840 VSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTE 899
Db        856 GHGALKSTTWNINHPIFPANSEVTFNVTFDVDSHASFGNKLLLKAIVASENNMSRTHKTK 915

Qy        900 FQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        916 FQLELPVKYAIYMIVTSDESSIRYLNFTASEMTSKVIQHQYQFNNLGQRSLPVSVVFWIP 975

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        976 VQINNVTVWDHPQVIFSQNLSSACHTEQKSPPHSNFRDQLERTPVLNCSVAVCKRIQCDL 1035

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1036 PSFNTQEIFNVTLKGNLSFDWYIKTSHGHLLLVSSTEILFNDSAFALLPGQESYVRSKTE 1095

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1096 TKVEPYEVHNPVPLIVGSSIGGLVLLALITAGLYKLGFFKRQYKDMMNEAAPQDAPPQ 1153
```

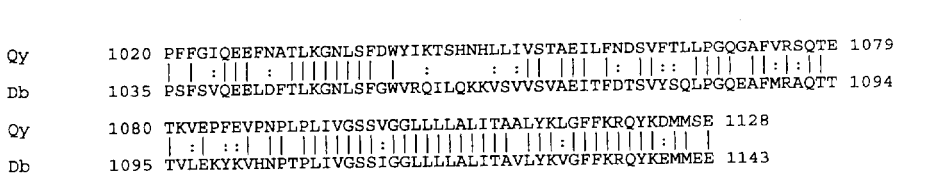
RESULT 3
RWHU1C
cell surface glycoprotein CD11c precursor - human
N;Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A36584; A35543; S00864
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A;Reference number: A36584
A;Contents: erratum
A;Accession: A36584
A;Molecule type: DNA
A;Residues: 1-1163 <COR>
A;Cross-references: UNIPROT:P20702
A;Note: this revision to the sequence from reference A35543 includes the carboxyl end
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.
A;Reference number: A35543; MUID:90153906; PMID:2303426
A;Accession: A35543
A;Molecule type: DNA
A;Residues: 1-834 <CO2>
A;Note: this sequence has been revised in reference A36584
R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A;Reference number: S00864; MUID:88166645; PMID:3327687
A;Accession: S00864
A;Molecule type: mRNA
A;Residues: 1-755,'L',757-1163 <CO3>
A;Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830
A;Note: part of this sequence was confirmed by protein sequencing
C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my
C;Genetics:
A;Gene: GDB:ITGAX; CD11C
A;Cross-references: GDB:119758; OMIM:151510
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C;Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F;20-1107/Domain: extracellular #status predicted <EXT>
F;149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F;1108-1133/Domain: transmembrane #status predicted <TMM>
F;1134-1163/Domain: intracellular #status predicted <INT>
F;61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

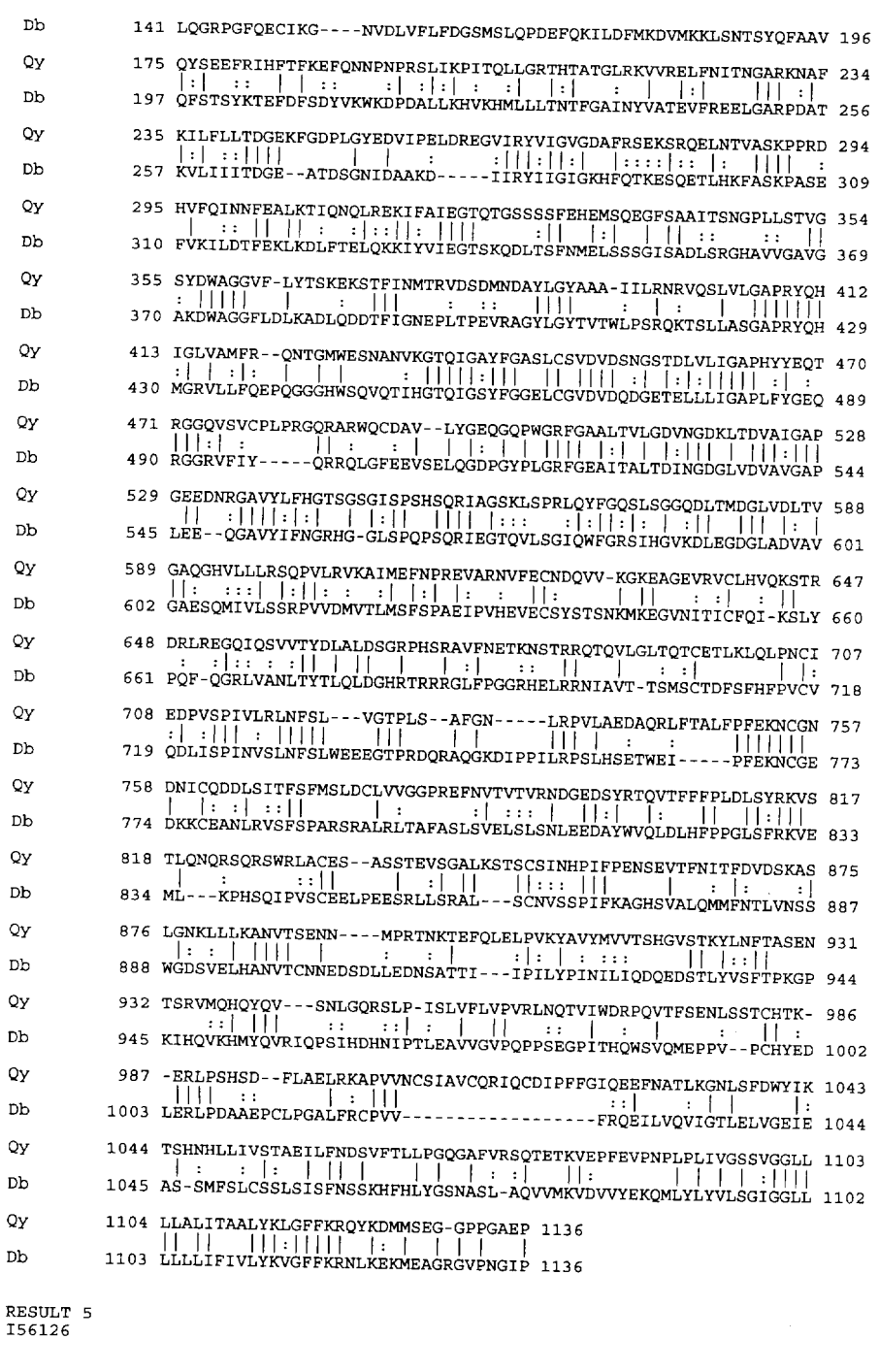
Query Match 59.0%; Score 3469; DB 1; Length 1163;
Best Local Similarity 60.9%; Pred. No. 3.1e-233;
Matches 688; Conservative 142; Mismatches 293; Indels 6; Gaps 4;



RESULT 4
S03308
cell surface glycoprotein CD11a precursor - human
N;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte functior
C;Species: Homo sapiens (man)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: S03308; A47458; A47565; A48759; S36044
R;Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.
J. Cell Biol. 108, 703-712, 1989
A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha subuni
A;Reference number: S03308; MUID:89139587; PMID:2537322
A;Accession: S03308
A;Molecule type: mRNA
A;Residues: 1-1170 <LAR>
A;Cross-references: UNIPROT:P20701; UNIPROT:Q9UBC8; EMBL:Y00796; NID:g31421; PIDN:CAA68
A;Note: part of this sequence was confirmed by protein sequencing
R;Cornwell, R.D.; Gollahon, K.A.; Hickstein, D.D.
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993
A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a) pr
A;Reference number: A47458; MUID:93248261; PMID:8097887
A;Accession: A47458
A;Molecule type: DNA
A;Residues: 1-20 <COR>
A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBIP:130863)
R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993
A;Title: Identification of cell-specific and developmentally regulated nuclear factors
A;Reference number: A47565; MUID:93281759; PMID:8099450
A;Accession: A47565
A;Molecule type: DNA
A;Residues: 1-20 <SHE>
A;Cross-references: GB:M95609
R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.
J. Biol. Chem. 268, 19305-19311, 1993
A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter.
A;Reference number: A48759; MUID:93374910; PMID:8103515
A;Accession: A48759
A;Molecule type: DNA
A;Residues: 1-20 <NUE>
A;Cross-references: EMBL:Z22804; NID:g311405; PIDN:CAA80461.1; PID:g311406
C;Genetics:
A;Gene: GDB:ITGAL; CD11A
A;Cross-references: GDB:119757; OMIM:153370
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen; tr
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted
F;154-317/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 26.2%; Score 1543.5; DB 2; Length 1170;
Best Local Similarity 34.3%; Pred. No. 6.2e-99;
Matches 403; Conservative 210; Mismatches 460; Indels 101; Gaps 37;

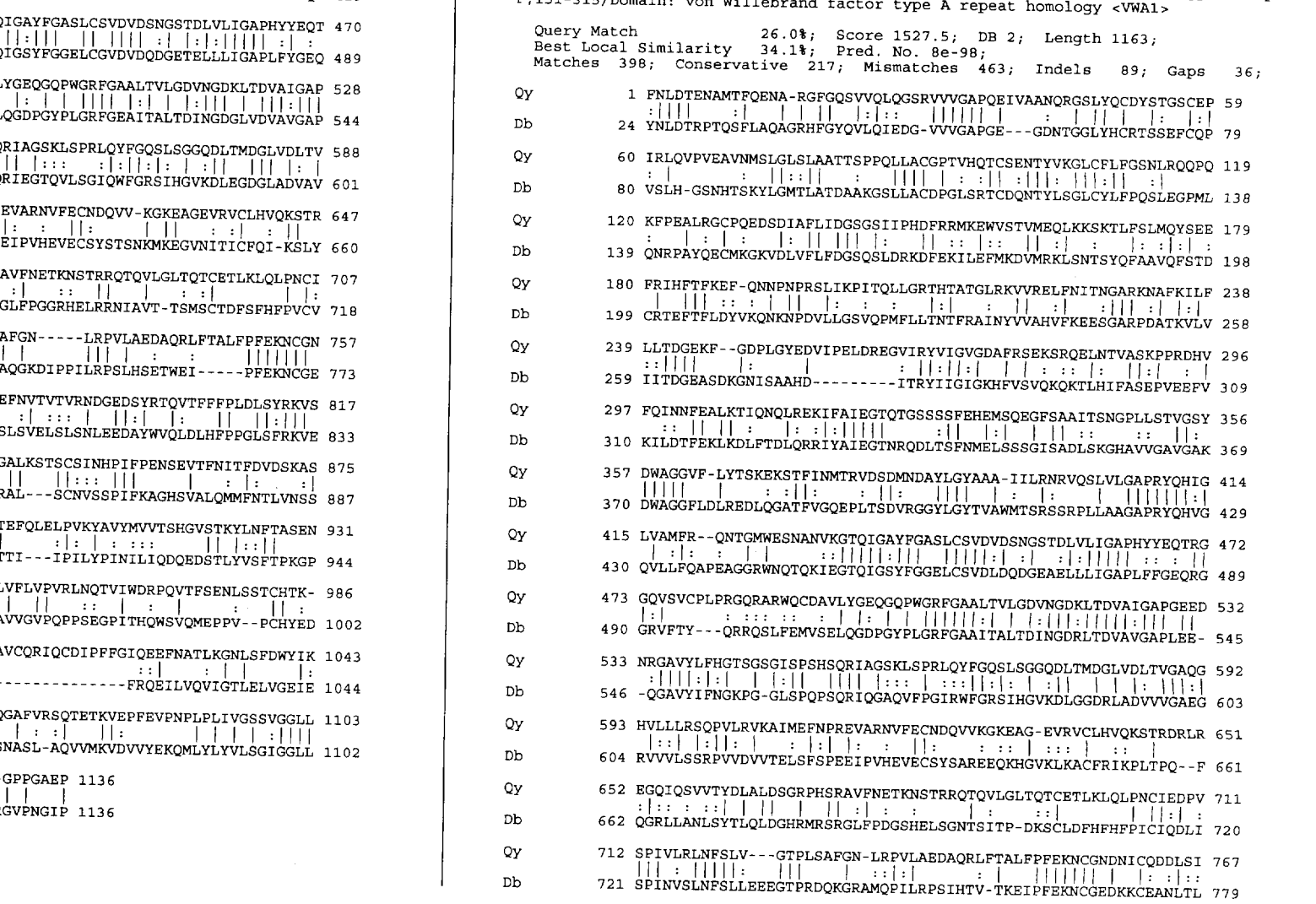


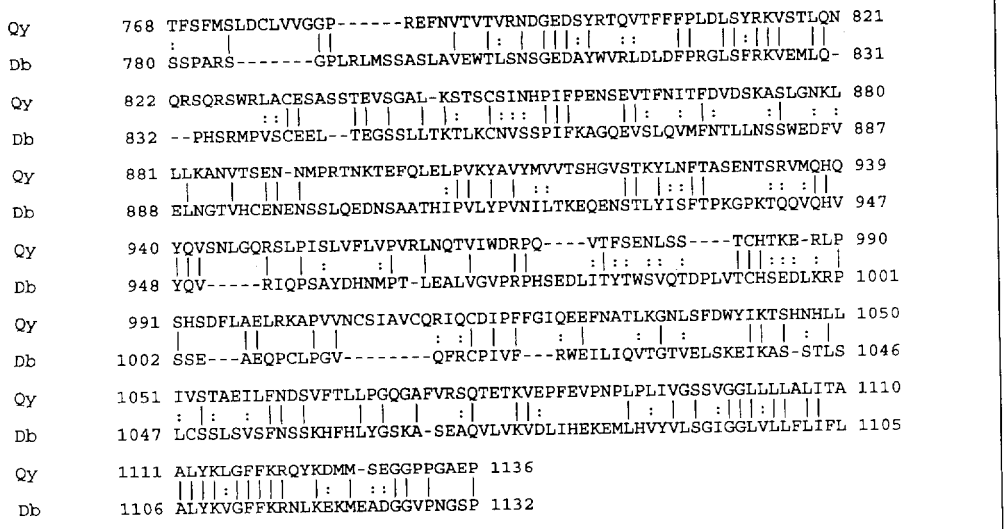


RESULT 5
I56126

lymphocyte fuction-associated molecule-1-alpha - mouse
C;Species: Mus musculus (house mouse)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: I56126
R;Kaufmann, Y.; Tseng, E.; Springer, T.A.
J. Immunol. 147, 369-374, 1991
A;Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit
A;Reference number: I56126; MUID:91268576; PMID:2051027
A;Accession: I56126
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1163 <RES>
A;Cross-references: UNIPROT:P24063; GB:M60778; NID:g198785; PIDN:AAA39426.1; PID:g19878
C;Genetics:
A;Gene: LFA-1
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
F;151-315/Domain: von Willebrand factor type A repeat homology <VWA1>

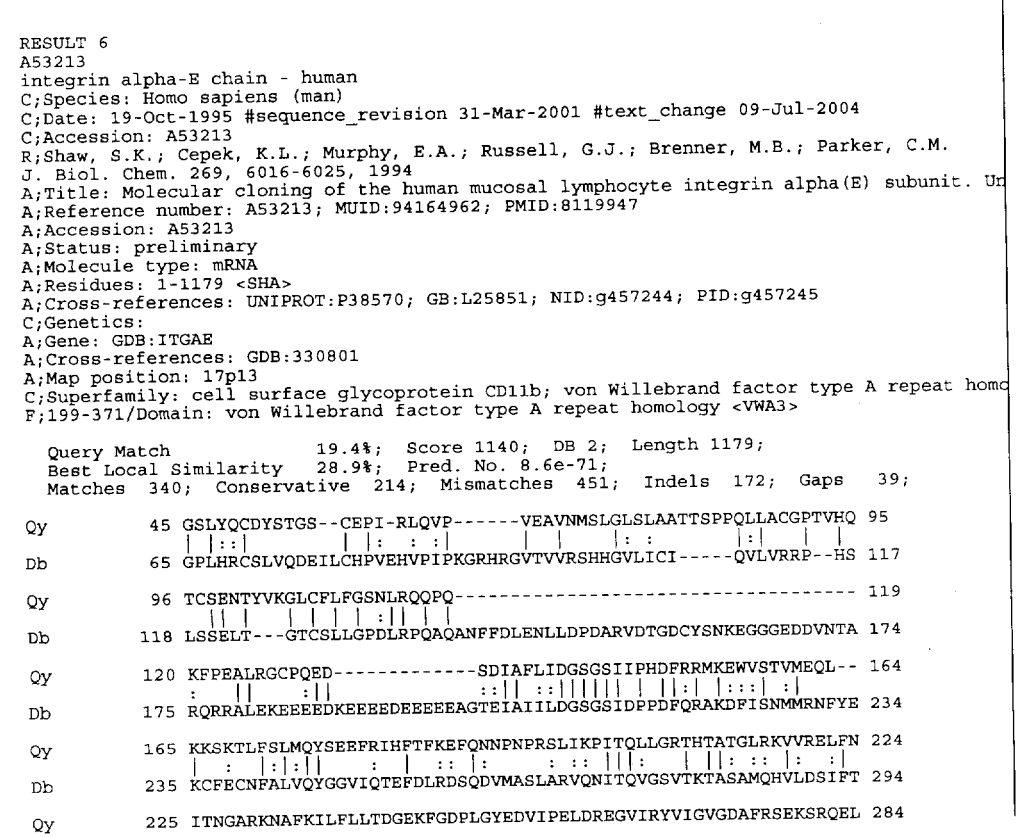
Query Match 26.0%; Score 1527.5; DB 2; Length 1163;
Best Local Similarity 34.1%; Pred. No. 8e-98;
Matches 398; Conservative 217; Mismatches 463; Indels 89; Gaps 36;

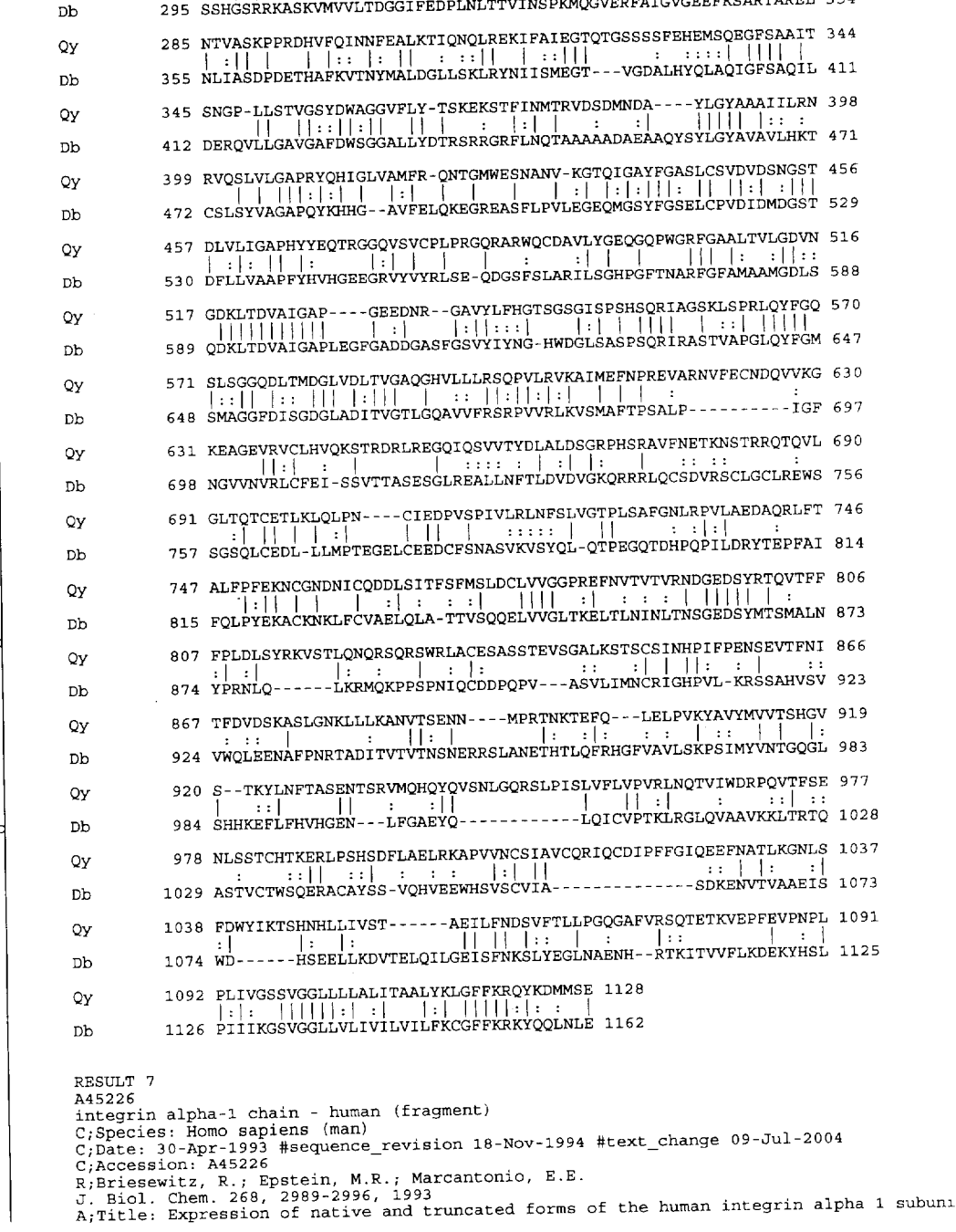




RESULT 6
A53213
integrin alpha-E chain - human
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 09-Jul-2004
C;Accession: A53213
R;Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6016-6025, 1994
A;Title: Molecular cloning of the human mucosal lymphocyte integrin alpha(E) subunit. Un
A;Reference number: A53213; MUID:94164962; PMID:8119947
A;Accession: A53213
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1179 <SHA>
A;Cross-references: UNIPROT:P38570; GB:L25851; NID:g457244; PID:g457245
C;Genetics:
A;Gene: GDB:ITGAE
A;Cross-references: GDB:330801
A;Map position: 17p13
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
F;199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 19.4%; Score 1140; DB 2; Length 1179;
Best Local Similarity 28.9%; Pred. No. 8.6e-71;
Matches 340; Conservative 214; Mismatches 451; Indels 172; Gaps 39;





RESULT 7
A45226
integrin alpha-1 chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A45226
R;Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.
J. Biol. Chem. 268, 2989-2996, 1993
A;Title: Expression of native and truncated forms of the human integrin alpha 1 subuni

A;Reference number: A45226; MUID:93155124; PMID:8428973
A;Accession: A45226
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1151 <BRI>
A;Cross-references: UNIPROT:P56199
A;Experimental source: hepatoblastoma cell line HepG2
A;Note: sequence extracted from NCBI backbone (NCBIP:124326)
F;142-317/Domain: von Willebrand factor type A repeat homology <VWA1>

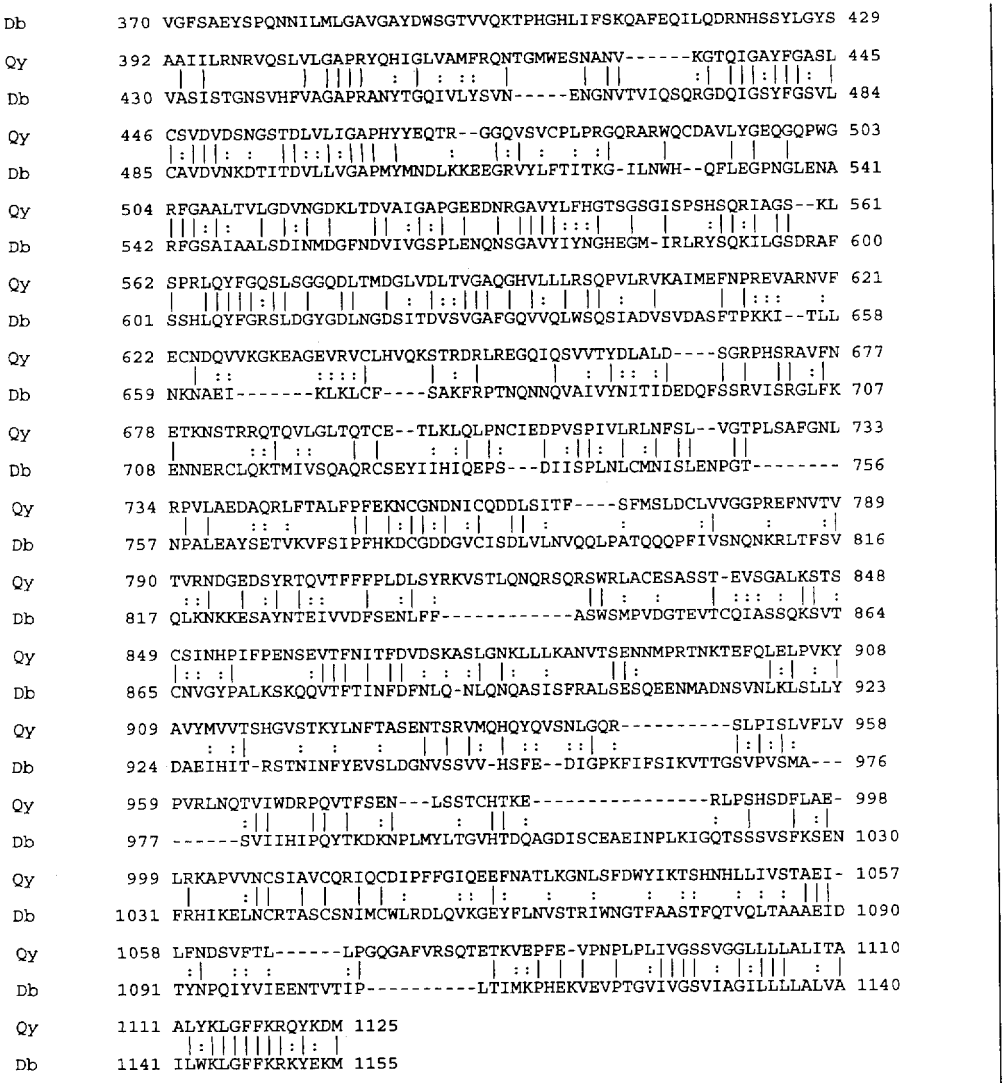
Query Match 18.6%; Score 1093.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 1.5e-67;
Matches 345; Conservative 206; Mismatches 487; Indels 203; Gaps 44;

```
Qy          1 FNLDTENAMTFQENARG-FGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQCDYSTGS 56
              ||:| :|:|||      || :| |     :|  |::|:|       |: | :|:|    |
Db          1 FNVDVKNSMTFSGPVEDMFGYTVQQYENEEGKWVLIGSPLVGQPKNRTGDVYKCPVGRGE 60

Qy         57 CEP-IRLQVPVEA---------VNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGL 106
               | ::| :||           ||: | :|  |       |||||      |     |  |:
Db         61 SLPCVKLDLPVNTSIPNVTEVEKNMTFGSTL-VTNPNGGFLACGPLYAYRCGHLHYTTGI 119

Qy        107 CFLFGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKK 166
              |        :      |  :: |  :  ||  ::||| || | |       |:  :  | |
Db        120 CSDVSPTFQVVNSIAP--VQECSTQ-LDIVIVLDGSNSIYPWDS------VTAFLNDLLK 170

Qy        167 ------SKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLGR-THTATGLRKVV 219
                   :|   ::|| |      |   :: :       | | |  || | || |
Db        171 RMDIGPKQTQVGIVQYGENVTHEFNLNKYSSTEEVLVAAKKIVQRGGRQTMTALGTDTAR 230

Qy        220 RELFNITNGARKNAFKILFLLTDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFR--- 276
              :| |    |||:   |:: ::|||| |    : || : : | :|: | :  ::
Db        231 KEAFTEARGARRGVKKVMVIVTDGESH-DNHRLKKVIQDCEDENIQRFSIAILGSYNRGN 289

Qy        277 --SEKSRQELNTVASKPPRDHVFQINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEM 334
                 :||  :|: ::||:|    | | :::  || ||    | |:|||:| |   |::||| ||
Db        290 LSTEKFVEEIKSIASEPTEKHFFNVSDELALVTIVKTLGERIFALEATADQSAASFEMEM 349

Qy        335 SQEGFSAAITSNGPLLSTVGSYDWAGGVFLYTSKE-----KSTF-INMTRVDSDMNDAYL 388
              || ||||  : :  :|  ||:||| | | :  : :       :|| :  |: :  :  :||
Db        350 SQTGFSAHYSQDWVMLGAVGAYDWNGTVVMQKASQIIIPRNTTFNVESTKKNEPL-ASYL 408

Qy        389 GYAAAIILRNRVQSL-VLGAPRYQHIGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCS 447
              ||       :   | : | ||| | | | ::|   |  :   : | |||:|||: | :
Db        409 GYTVNSATASSGDVLYIAGQPRYNHTGQVIIYRMEDGNIKILQTLSGEQIGSYFGSILTT 468

Qy        448 VDVDSNGSTDLVLIGAPHY--------------YEQTR-GGQVSVCPLPRGQRARWQCD 491
              |:| : :||::|:||| |                 |||   |:|: |: :  :  | :
Db        469 TDIDKDSNTDILLVGAPMYMGTEKEEQGKVYVYALNQTRFEYQMSLEPIKQTCCSSRQHN 528

Qy        492 AVLYGEQGQPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISP 550
              :    : :| | ||| |:  : |:| |   |: |||| |:|: ||||::|| ||  |
Db        529 SCTTENKNEPCGARFGTAIAAVKDLNLDGFNDIVIGAPLEDDHGGAVYIYHG-SGKTIRK 587

Qy        551 SHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIME 610
              ::|||       |::||||: |  ||  ||| |:|:|  |    |  |: |  ||  |
Db        588 EYAQRIPSGGDGKTLKFFGQSIHGEMDLNGDGLTDVTIGGLGGAALFWSRDVAVVKVTMN 647

Qy        611 FNPREVARNVFECNDQVVKGKEAG--EVRVCLHVQ-KSTRDRLREGQIQSVVTYDLALDS 667
              | | :|      |:   ::|||       ||  |: ||  | : |  :|    | : |||
Db        648 FEPNKVNIQKKNCH---MEGKETVCINATVCFEVKLKSKEDTIYEADLQ----YRVTLDS 700

Qy        668 GRPHSRAVFNET------KNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNFS 721
              |  ||: |: |       :| | |:::      |  :  :   : :| |    : |:|:
Db        701 LRQISRSFFSGTQERKVQRNITVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDFN 752

Qy        722 LVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVGG 781
              |   | :      |||  :          || |:|||    | |||:  :     | |:|
Db        753 LT-DPENG-----PVLDDSLPNSVHEYIPFAKDCGNKEKCISDLSLHVATTEKDLLIVRS 806

Qy        782 PRE-FNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              : |||::||:|  : :| |:    :  :| :  :  :|          :|||
Db        807 QNDKFNVSLTVKNTKDSAYNTRTIVHYSPNLVFSGIEAIQKD--------SCESN----- 853

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGN-KLLLKANVTSENNMPRTNKTE 899
                    : :| : :|        ||| | |  ::   : |  : | |   ||      :
Db        854 ----HNITCKVGYPFLRRGEMVTFKILFQFNTSYLMENVTIYLSATSDSEEPPETLSDNV 909

Qy        900 FQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSN------LGQRS---- 949
                : :|||| | :   |  |  :::  |:|    |:      : |      | ::|
Db        910 VNISIPVKYEVGLQFYS-SASEYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSFP 968

Qy        950 ---LPISLVF----------LVPVRLNQTVIWDRPQVTFSENLSSTCHTKE--------- 987
                  | :|: |          | |  |:            |||  :   |  |
Db        969 MPELKLSISFPNMTSNGYPVLYPTGLSS----------SENANCRPHIFEDPFSINSGK 1017

Qy        988 RLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPPFGIQE----------------EFN 1029
              :: : :|    |::  :::|:       | |::    | :                 |
Db       1018 KMTTSTD---HLKRGTILDCNTCKFATITCNLTSSDISQVNVSLILWKPTFIKSYFSSLN 1074

Qy       1030 ATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPN 1089
              |::| |          | |  |::|::                  :  :   ::    :|
Db       1075 LTIRGEL-------RSENASLVLSSSN----------------QKRELAIQISKDGLPG 1110

Qy       1090 PLPL--IVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
                :||  |: |:  ||||| |:  ||:|:|||||  |  | :
Db       1111 RVPLWVILLSAFAGLLLLMLLILALWKIGFFKRPLKKKMEK 1151
```

RESULT 8
I45914
integrin alpha 2 subunit - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C;Accession: I45914
R;Kamata, T.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A;Title: Identification of putative ligand binding sites within the I-domain of integri
A;Reference number: A54402; MUID:94193647; PMID:7511592
A;Accession: I45914
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1170 <KAM>
A;Cross-references: UNIPROT:P53710; GB:L25886; NID:g439695; PIDN:AAB59255.1; PID:g43969
F;161-336/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.2%; Score 1071; DB 2; Length 1170;
Best Local Similarity 27.4%; Pred. No. 5.5e-66;
Matches 333; Conservative 219; Mismatches 495; Indels 168; Gaps 47;

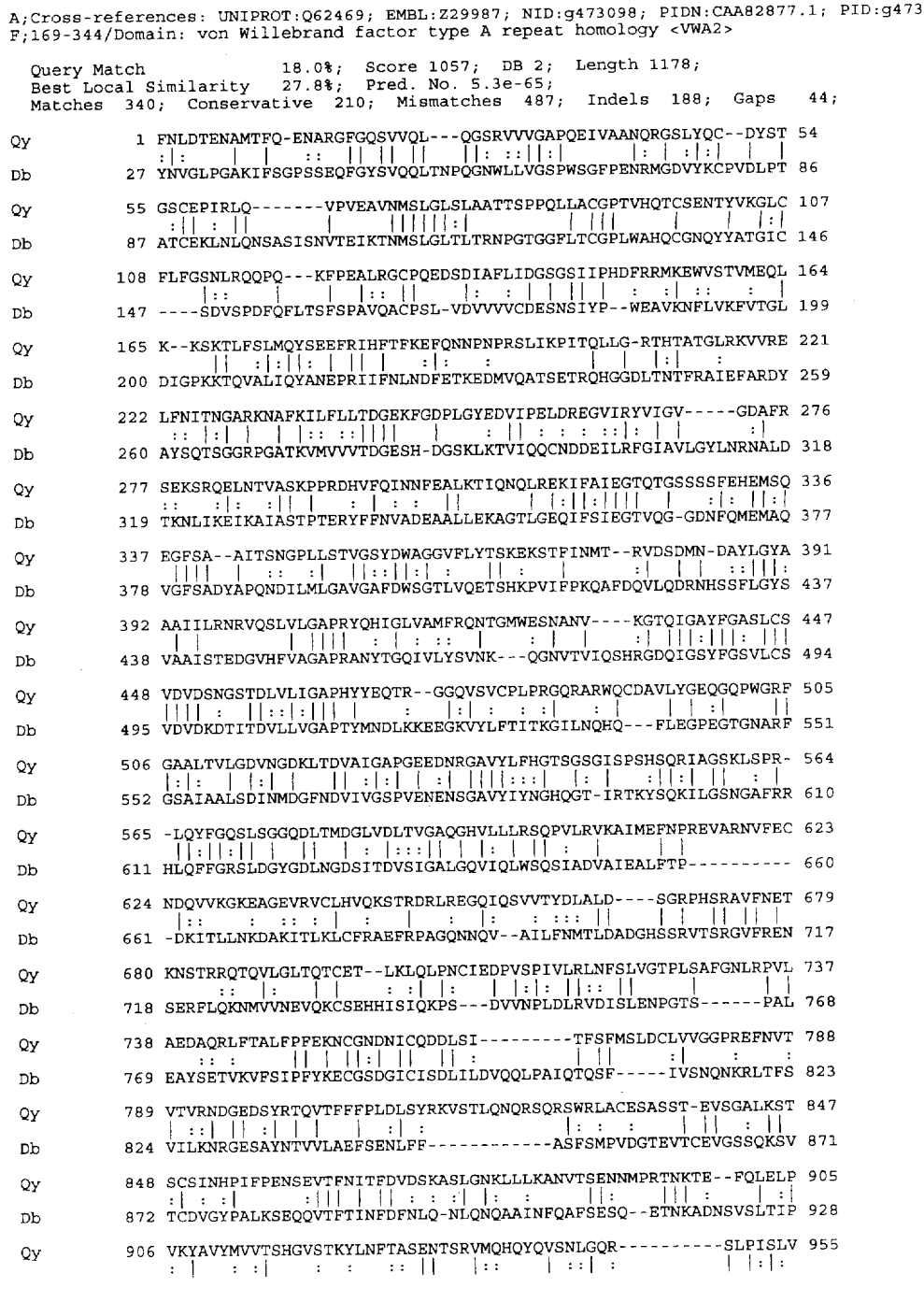
```
Qy          1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQC--DYST 54
              :|:    |  |   ::  || :| |    :|: ::||:|      |: | :|:|  | ||
Db         19 YNVGLPKAKIFSGPSSEQFGYAVQQFINPKGNWLLVGSPWSGFPKNRMGDVYKCPVDLST 78

Qy         55 GSCEPIRLQ-------VPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLC 107
               :|| : ||       |     ||||||:|       | |||   | |    | |:|
Db         79 TTCEKLNLQTSTSMSNVTEMKTNMSLGLTLTRNVGTGGFLTCGPLWAQQCGSQYYTTGVC 138

Qy        108 FLFGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLK-- 165
                  :  |    |  |:: ||     |:  : | | || | |   :| ::   :: |
Db        139 SDVSPDF-QLRTSFAPAVQTCP-SFIDVVVVCDESNSIYPWD--AVKNFLEKFVQGLDIG 194

Qy        166 KSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLIKPITQLL----GRTHTATGLRKVVRE 221
                :||   |:||:   |: |    |::      :||  :|      |:|   ::
Db        195 PTKTQMGLIQYANNPRVVFNLNTFKSKD---EMIKATSQTFQYGGDLTNTFKAIQYARDT 251

Qy        222 LFNITNGARKNAFKILFLLTDGEKFGDPLGYEDVIPELDREGVIRYVIGV-----GDAFR 276
                ::   | |  | |:: ::|||   |    : || : ::: ::|: | |       :|
Db        252 AYSTAAGGRPGATKVMVVVTDGESH-DGSKLKAVIDQCNKDNILRFGIAVLGYLNRNALD 310

Qy        277 SEKSRQELNTVASKPPRDHVFQINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQ 336
                ::   :|:  :|| |  | | ::: |       : |:||:|||| |   :|: ||||
Db        311 TKNLIKEIKAIASIPTERHFFNVSDEADLLEKAGTIGEQIFSIEGTVQG-GDNFQMEMSQ 369

Qy        337 EGFSAAIT--SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYA 391
              ||||  :  :|  :|  ||:|||:| |   |      |      ::  | |  :||||:
```



RESULT 9
S44142
VLA-2 protein homolog - mouse
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S44142
R;Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze
submitted to the EMBL Data Library, January 1994
A;Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not
A;Reference number: S44142
A;Accession: S44142
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1178 <EDE>

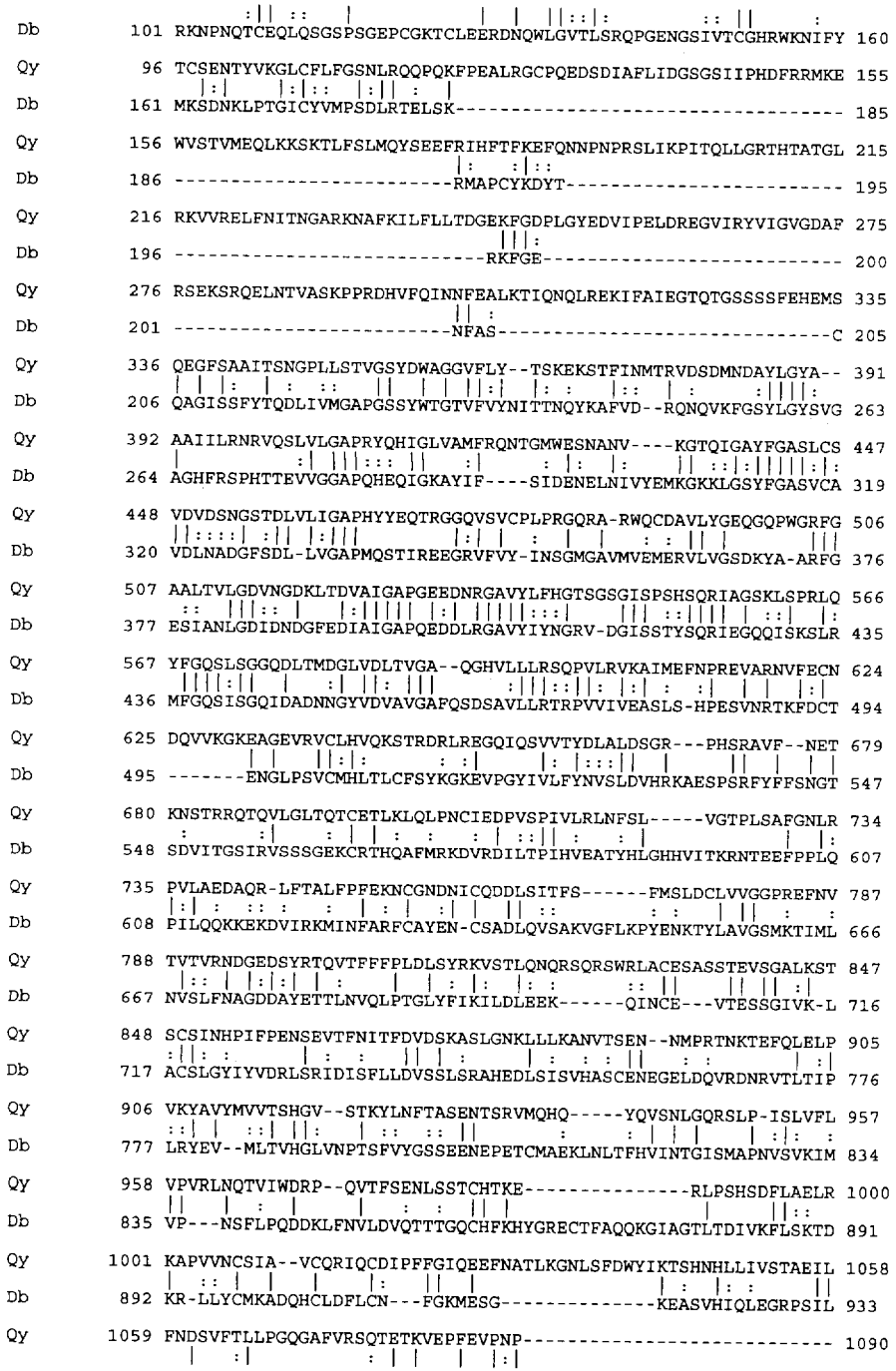
A;Cross-references: UNIPROT:Q62469; EMBL:Z29987; NID:g473098; PIDN:CAA82877.1; PID:g473
F;169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

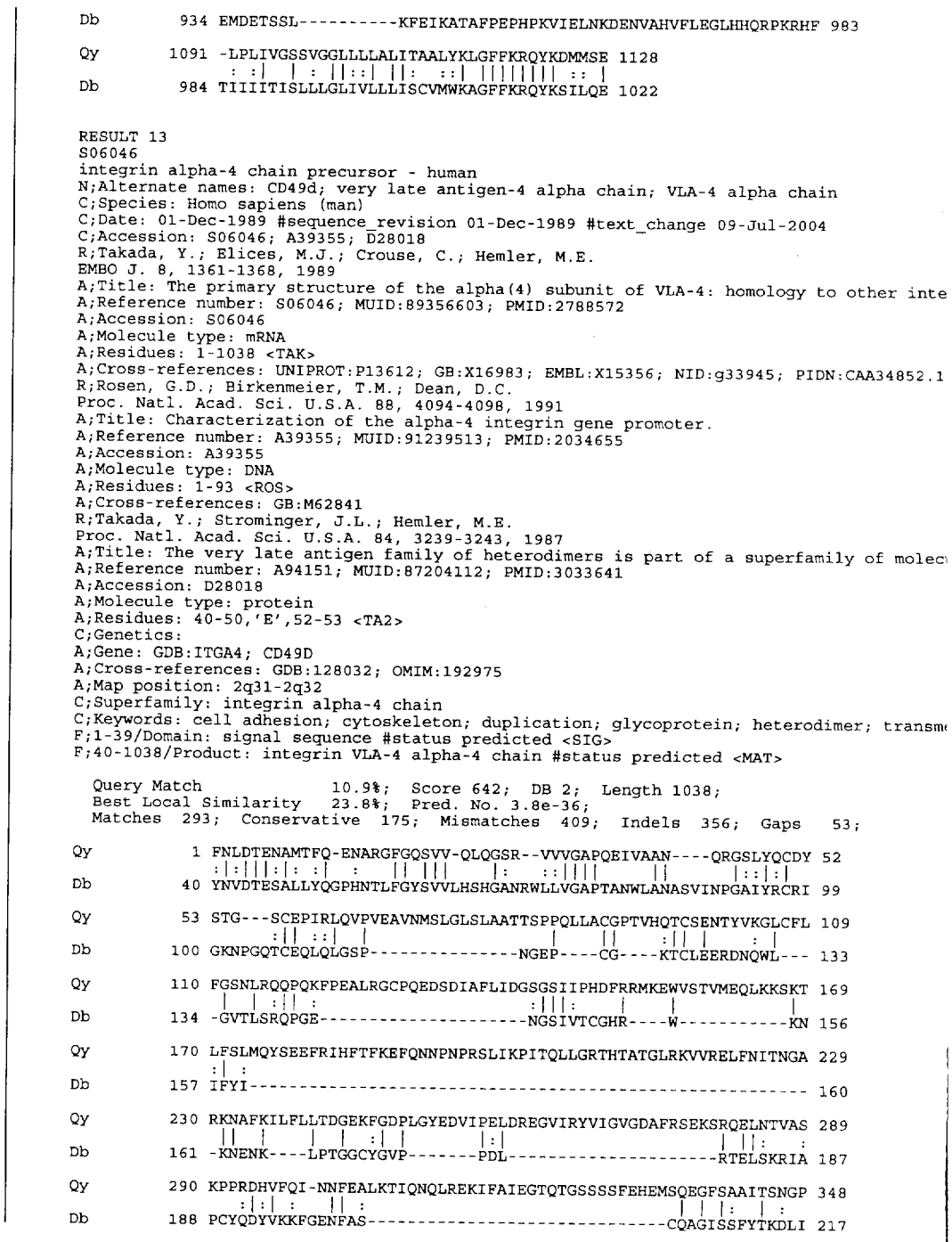
Query Match 18.0%; Score 1057; DB 2; Length 1178;
Best Local Similarity 27.8%; Pred. No. 5.3e-65;
Matches 340; Conservative 210; Mismatches 487; Indels 188; Gaps 44;



```
Db        929 LLYDAELHLT-RSTNINFYEISSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVSMA 984

Qy        956 FLV---------------------------------PVRLNQTVIWDRPQVTF-SENLS 980
                :                                 |::|  |     | |:| :||
Db        985 LVTIHIPQYTKEKNPLLYLTGIQTDQAGDISCTAEINPLKLPHTA----PSVSFKNENFR 1040

Qy        981 STCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDW 1040
                ||||                ::|    |  | | :    :: |:   :   :
Db       1041 ---HTKE----------------LDCRTTSCSNITCWLKDLHMKAEYFINVTTRVWNRT 1080

Qy       1041 YIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVG 1100
              :  ::   : : : |||  ::    ::       :       |  |||    :|:|| :
Db       1081 FAASTFQTVQLTAAAEIDTHNPQLFVIEENAVTIPLMIMKPTEKAEVPT--GVIIGSIIA 1138

Qy       1101 GLLLLALITAALYKLGFFKRQYKDM 1125
              |:|||  :|| |:||||||||||| |
Db       1139 GILLLLAMTAGLWKLGFFKRQYKKM 1163


RESULT 10
A35854
integrin alpha-1 chain precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 09-Jul-2004
C;Accession: A35854; S11243
R;Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto,
J. Cell Biol. 111, 709-720, 1990
A;Title: Molecular cloning of the rat integrin alpha-1-subunit: a receptor for laminin a
A;Reference number: A35854; MUID:90338125; PMID:2380249
A;Accession: A35854
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1180 <IGN>
A;Cross-references: UNIPROT:P18614; GB:X52140; NID:g56493; PIDN:CAA36384.1; PID:g56494
C;Keywords: cell adhesion; cytoskeleton; transmembrane protein
F;170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

  Query Match             17.9%;  Score 1055;  DB 2;  Length 1180;
  Best Local Similarity   27.4%;  Pred. No. 7.3e-65;
  Matches  346;  Conservative  196;  Mismatches  475;  Indels  246;  Gaps   48;

Qy          1 FNLDTENAMTFQENARG-FGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQCDYSTGS 56
              ||:| :|:|:|      || :| |    :|  |::|:|       : | :|:|
Db         29 FNVDVKNSMSFSGPVEDMFGYTVQQYENEEGKWVLIGSPLVGQPKARTGDVYKCPVGRER 88

Qy         57 CEP-IRLQVPVEA---------VNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGL 106
               | ::| :||            ||: | :|  |      |||||       |    | |:
Db         89 AMPCVKLDLPVNTSIPNVTEIKENMTFGSTL-VTNPNGGFLACGPLYAYRCGHLHYTTGI 147

Qy        107 CFLFGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVM---EQ 163
              |       :      | :: |  :  ||  ::||| || |      | |:   :
Db        148 CSDVSPTFQVVNSFAP--VQECSTQ-LDIVIVLDGSNSIYP--------WESVIAFLNDL 196

Qy        164 LKK-----SKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLG-RTHTATGLRK 217
              ||:      :|  ::|| |     |  :: :         | :  | :| || |:
Db        197 LKRMDIGPKQTQVGIVQYGENVTHEFNLNKYSSTEEVLVAANKIGRQGGLQTMTALGIDT 256

Qy        218 VVRELFNITNGARKNAFKILFLLTDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFR- 276
              :| |    |||:   |:: ::|||    |    : || : : | : |: | :   :
Db        257 ARKEAFTEARGARRGVKKVMVIVTDGESH-DNYRLKQVIQDCEDENIQRFSIAILGHYNR 315

Qy        277 ----SEKSRQELNTVASKPPRDHVFQINNPEALKTIQNQLREKIFAIEGTQTGSSSSFEH 332
                  :||  :|: ::||:|    | | :::  || ||    | |:|||:| |   |::|||
Db        316 GNLSTEKFVEEIKSIASEPTEKHFFNVSDELALVTIVKALGERIFALEATADQSAASFEM 375

Qy        333 EMSQEGFSAAITSNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMND---AY 387
              |||| ||||  : :  :|  ||:||| | | :  : :      | |      : ||:   :|
Db        376 EMSQTGFSAHYSQDWVMLGAVGAYDWNGTVVMQKANQMVIPHNTTFQTEPAKMNEPLASY 435

Qy        388 LGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCS 447
              |||              : | ||| | | | :::   |       : | |||:|||: | :
Db        436 LGYTVNSATIPGDVLYIAGQPRYNHTGQVVIYKMEDGNINILQTLGGEQIGSYFGSVLTT 495

Qy        448 VDVDSNGSTDLVLIGAPHY----------YEQTR-GGQVSVCPLPRGQRARWQCD 491
              :|:| :  |||:|:|||  |                |||   |:|: |: :   :  : :
Db        496 IDIDKDSYTDLLLVGAPMYMGTEKEEQGKVYVYAVNQTRFEYQMSLEPIRQTCCSSLKDN 555

Qy        492 AVLYGEQGQPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISP 550
              :     : :| | ||| |:  : |:| |   || |||| |:|: ||||::|| ||  |
Db        556 SCTKENKNEPCGARFGTAIAAVKDLNVDGFNDVVIGAPLEDDHAGAVYIYHG-SGKTIRE 614

Qy        551 SHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIME 610
              :::|||          |::|||:  |  ||  ||| |:|:|  |   |  :: |  || |
Db        615 AYAQRIPSGGDGKTLKFFGQSIHGEMDLNGDGLTDVTIGGLGGAALFWARDVAVVKVTMN 674

Qy        611 FNPREVARNVFECNDQVVKGKEAG--EVRVCLHVQ-KSTRDRLREGQIQSVVTYDLALDS 667
              | | :|      |   |:|||       :| ||: ||  | : |  :|    | : |||
Db        675 FEPNKVNIQKKNCR---VEGKETVCINATMCFHVKLKSKEDSIYEADLQ----YRVTLDS 727

Qy        668 GRPHSRAVFNET------KNSTRRQTQVLGLTQTCETLKLQLPNCI---------EDPVS 712
              |  ||: |: |       :| | |:::                ||          |
Db        728 LRQISRSFFSGTQERKIQRNITVRESE----------------CIRHSFYMLDKHDFQD 770

Qy        713 PIVLRLNFSLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFM 772
              : : |:|:|    | :       ||| :       || |:|||    |  ||::  |
Db        771 SVRVTLDFNLT-DPENG-----PVLDDALPNSVHEHIPFAKDCGNKERCISDLTLNVSTT 824

Qy        773 SLDCLVVGGPRE-FNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLA 831
                |:|    : |||::||:| |: :| |:       :| :  :  :|          :
Db        825 EKSLLIVKSQHDKFNVSLTVKNKGDSAYNTRTVVQHSPNLIFSGIEEIQKD--------S 876

Qy        832 CESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLL-LKANVTSEN 890
              |||          :: :| : :|       |||  | |  ::      | :: | |   ||
Db        877 CESN---------QNITCRVGYPFLRAGETVTFKIIFQFNTSHLSENAIIHLSATSDSEE 927

Qy        891 NMPRTNKTEFQLELPVKYAV----YMVVTSHGVST-------KYLNFTASENTSRVMQHQ 939
               :   |  |  : :|||| |     |   : | :|         :::| |  |:    :
Db        928 PLESLNDNEVNISIPVKYEVGLQFYSSASEHHISVAANETIPEFINST--EDIGNEINVF 985

Qy        940 YQVSNLGQRSLP---ISLVF----------LVPVRLNQTVIWD-------RP-------- 971
              | :   |  :|   :|: |          | |:       |          ||
Db        986 YTIRKRGHFPMPELQLSISFPNLTADGYPVLYPIG------WSSSDNVNCRPRSLEDPFG 1039

Qy        972 -----QVTFS----------ENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQ 1016
                  ::| |           :: ||||                : ||:       |
Db       1040 INSGKKMTISKSEVLKRGTIQDCSSTC----------------GVATITCSLLPSDLSQ 1082

Qy       1017 CDI------PFFGIQEEF---NATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLL 1067
               ::       | | |:  |   | ||:| |          | |  | :|::
Db       1083 VNVSLLLWKPTF-IRAHFSSLNLTLRGELK-------SENSSLTLSSSN----------- 1123

Qy       1068 PGQGAFVRSQTETKVEPFEVPNPLPL--IVGSSVGGLLLLALITAALYKLGFFKRQYKDM 1125
                      : :   ::    :|  :|| |: |:  ||||| |:  ||:|:|||||  |
Db       1124 ------RKRELAIQISKDGLPGRVPLWVILLSAFAGLLLLMLLILALWKIGFFKRPLKKK 1177

Qy       1126 MSE 1128
              | :
Db       1178 MEK 1180


RESULT 11
A33998
integrin alpha-2 chain precursor - human
N;Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 09-Jul-2004
C;Accession: A33998; B56793; A53117
R;Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet
A;Reference number: A33998; MUID:89308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: UNIPROT:P17301; GB:X17033; NID:g33906; PIDN:CAA34894.1; PID:g33907
A;Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
```

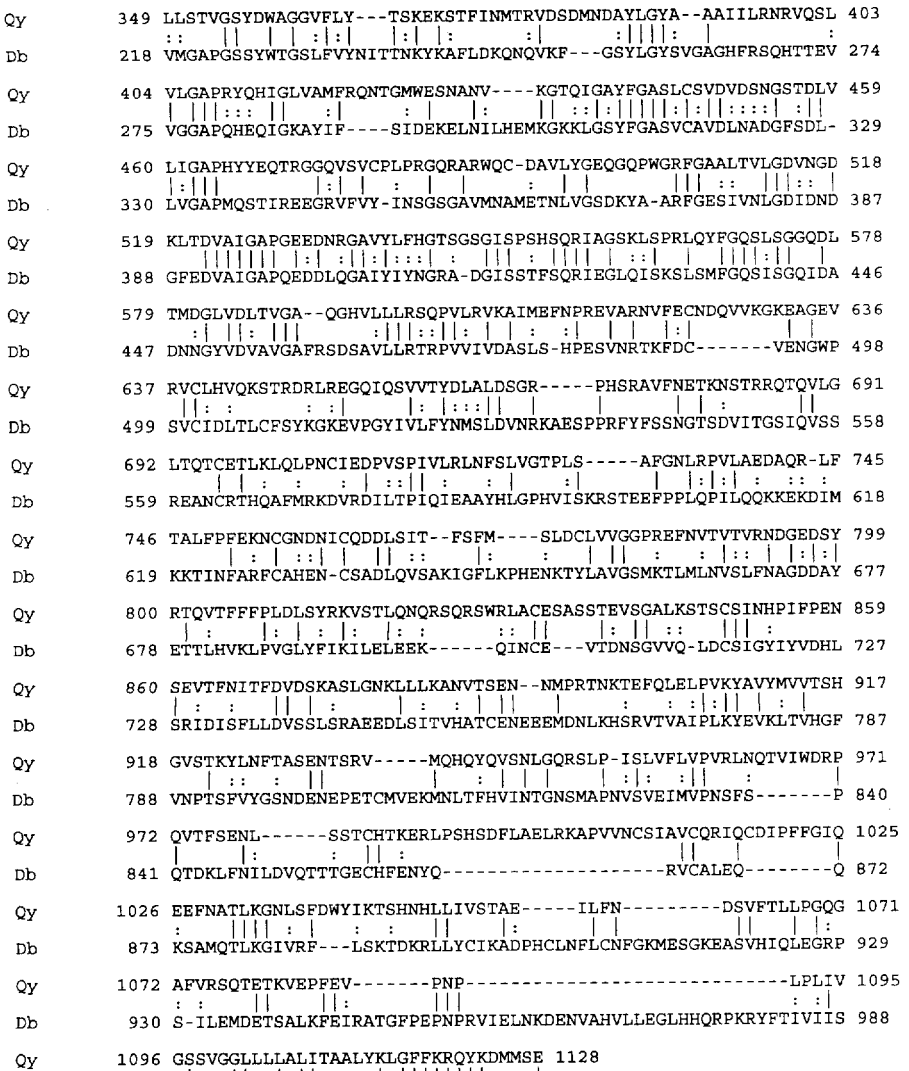
```
R;Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIc, GPIc*, GPIIa ar
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16,'V',18-21 <ZUT>
A;Cross-references: GB:L24121; NID:g400342; PIDN:AAA16619.2; PID:g4583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TMM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

  Query Match             17.9%;  Score 1054;  DB 2;  Length 1181;
  Best Local Similarity   26.7%;  Pred. No. 8.6e-65;
  Matches  327;  Conservative  216;  Mismatches  494;  Indels  188;  Gaps   43;

Qy          1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQC--DYST 54
              :|:    |  |   ::  || :| |     :|: ::||:|       |: | :|:|  | ||
Db         30 YNVGLPEAKIFSGPSSEQFGYAVQQFINPKGNWLLVGSPWSGFPENRMGDVYKCPVDLST 89

Qy         55 GSCEPIRLQ-------VPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLC 107
               :|| : ||       |     |||||| |          | |||   | |     |  |:|
Db         90 ATCEKLNLQTSTSIPNVTEMKTNMSLGLILTRNMGTGGFLTCGPLWAQQCGNQYYTTGVC 149

Qy        108 FLFGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLK-- 165
                  :  |    |  | : ||    |:  : | | || | |   :| ::   :: |
Db        150 SDISPDF-QLSASFSPATQPCPSL-IDVVVVCDESNSIYPWD--AVKNFLEKFVQGLDIG 205

Qy        166 KSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLG-RTHTATGLRKVVRELFN 224
               :||   |:||:   |: |    ::             :|  |  |:|   ::   :  ::
Db        206 PTKTQVGLIQYANNPRVVFNLNTYKTKEEMIVATSQTSQYGGDLTNTFGAIQYARKYAYS 265

Qy        225 ITNGARKNAFKILFLLTDGEKFGDPLGYEDVIPELDREGVIRYVIGV-----GDAFRSEK 279
               :| |::| |:: ::||||   |    : || : : : ::|: | |      :|  ::
Db        266 AASGGRRSATKVMVVVTDGESH-DGSMLKAVIDQCNHDNILRFGIAVLGYLNRNALDTKN 324

Qy        280 SRQELNTVASKPPRDHVFQINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGF 339
               :|:  :|| |   : | ::::  ||       | |:||:|||| |   :|: |||| ||
Db        325 LIKEIKAIASIPTERYFFNVSDEAALLEKAGTLGEQIFSIEGTVQG-GDNFQMEMSQVGF 383

Qy        340 SAAITSNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYAAAI 394
              ||  :|     :|  ||:: |:| :   ||       |       ::  | |  :||||: |
Db        384 SADYSSQNDILMLGAVGAFGWSGTIVQKTSHGHLIFPKQAFDQILQDRNHSSYLGYSVAA 443

Qy        395 ILRNRVQSLVLGAPRYQHIGLVAMFRQNTGMWESNANV------KGTQIGAYFGASLCSV 448
              |     | ||||  :| : ::  |       | |:       :| |||:|||: ||||
Db        444 ISTGESTHFVAGAPRANYTGQIVLYSVN-----ENGNITVIQAHRGDQIGSYFGSVLCSV 498

Qy        449 DVDSNGSTDLVLIGAPHYYEQTR--GGQVSVCPLPRGQRARWQCDAVLYGEQGQPWGRFG 506
              ||| :  ||::|:||| |     :   |:| :  : :|    : |     | | :|    |||
Db        499 DVDKDTITDVLLVGAPMYMSDLKKEEGRVYLFTIKKGILGQHQ---FLEGPEGIENTRFG 555

Qy        507 AALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGS--KLSPR 564
              :|:  | |:| |    || :|:| |  | ||||:::|  |: |   :||:| ||
Db        556 SAIAALSDINMDGFNDVIVGSPLENQNSGAVYIYNGHQGT-IRTKYSQKILGSDGAFRSH 614

Qy        565 LQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECN 624
              |||||:|| |  ||  | : |:::|| | |: | || :  |       | | ::   :  |
Db        615 LQYFGRSLDGYGDLNGDSITDVSIGAFGQVVQLWSQSIADVAIEASFTPEKI--TLVNKN 672

Qy        625 DQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVTYDLALD----SGRPHSRAVFNETK 680
               |::      :::|    | : |  :   |  : |:: ||    | |  || :| |
Db        673 AQII-------LKLCF----SAKFRPTKQNNQVAIVYNITLDADGFSSRVTSRGLFKENN 721

Qy        681 NSTRRQTQVLGLTQTC--ETLKLQLPNCIEDPVSPIVLRLNFSLVGTPLSAFGNLRPVLA 738
                 ::  |:   |:|    : :| |:   | |: :|:: ||      |  | |
Db        722 ERCLQKNMVVNQAQSCPEHIIYIQEPS---DVVNSLDLRVDISLENPGTS------PALE 772

Qy        739 EDAQRLFTALFPFEKNCGNDNICQDDLSITF----SFMSLDCLVVGGPREFNVTVTVRND 794
               ::        || |:|| | :|  || :      :      :|   :    :||:|
Db        773 AYSETAKVFSIPFHKDCGEDGLCISDLVLDVRQIPAAQEQPFIVSNQNKRLTFSVTLKNK 832

Qy        795 GEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASST-EVSGALKSTSCSINH 853
               | :| | :   |  :| :              |: |  :    | :|: : || :| : :
Db        833 RESAYNTGIVVDFSENLFF------------ASFSLPVDGTEVTCQVAASQKSVACDVGY 880

Qy        854 PIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEFQLELPVKYAVYMV 913
              |      :||| | || : : :| |:  |      ||:           |::|: |   :
Db        881 PALKREQQVTFTINFDFNLQ-NLQNQASLSFQALSESQEENKADNLVNLKIPLLYDAEI- 938

Qy        914 VTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQR----------SLPISLVFLV----- 958
                |   :  :||    :   |    :   ::| :           |:|:|:  ::
Db        939 ---HLTRSTNINFYEISSDGNVPSIVHSFEDVGPKFIFSLKVTTGSVPVSMATVIIHIPQ 995

Qy        959 ----------------------------PVRLNQTVIWDRPQVTF-SENLSSTCHTKER 988
                                         |::: ||       |:| |||     ||||
Db        996 YTKEKNPLMYLTGVQTDKAGDISCNADINPLKIGQT----SSSVSFKSENFR---HTKE- 1047

Qy        989 LPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNH 1048
                          :||  | |  : | :    :: |:   :   :   :  ::
Db       1048 ---------------LNCRTASCSNVTCWLKDVHMKGEYFVNVTTRIWNGTFASSTFQT 1091

Qy       1049 LLIVSTAEI-LFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNP-----LP--LIVGSSVG 1100
              : : : |||  :|  :: :          :  |    |  :  |     :|  :|:|| :
Db       1092 VQLTAAAEINTYNPEIYVI----------EDNTVTIPLMIMKPDEKAEVPTGVIIGSIIA 1141

Qy       1101 GLLLLALITAALYKLGFFKRQYKDM 1125
              |:|||  :  | |:|||||||:|: |
Db       1142 GILLLLALVAILWKLGFFKRKYEKM 1166


RESULT 12
A41131
lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
N;Alternate names: integrin alpha-4
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004
C;Accession: A41131; S16742
R;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.L.
J. Cell Biol. 115, 1149-1158, 1991
A;Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyt
A;Reference number: A41131; MUID:92064645; PMID:1840602
A;Accession: A41131
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1039 <NEU>
A;Cross-references: UNIPROT:Q00651; EMBL:X53176; NID:g51484; PIDN:CAA37316.1; PID:g514
C;Superfamily: integrin alpha-4 chain
C;Keywords: cytoskeleton; transmembrane protein

  Query Match             11.3%;  Score 663;  DB 2;  Length 1039;
  Best Local Similarity   22.7%;  Pred. No. 1.3e-37;
  Matches  281;  Conservative  187;  Mismatches  403;  Indels  368;  Gaps   48;

Qy          1 FNLDTENAMTFQ-ENARGFGQSVV-QLQGSR--VVVGAPQEIVAAN----QRGSLYQCDY 52
              :||| |||: :|  :   || |||     ||:  ::||||      :|      |::|:|
Db         41 YNLDPENALLYQGPSGTLFGYSVVLHSHGSKRWLIVGAPTASWLSNASVVNPGAIYRCGI 100

Qy         53 STG---SCEPIRLQVP----------VEAVNMSLGLSLAATTSP-PQLLACG---PTVHQ 95
```





RESULT 13
S06046
integrin alpha-4 chain precursor - human
N;Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain
C;Species: Homo sapiens (man)
C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C;Accession: S06046; A39355; D28018
R;Takada, Y.; Elices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989
A;Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other inte
A;Reference number: S06046; MUID:89356603; PMID:2788572
A;Accession: S06046
A;Molecule type: mRNA
A;Residues: 1-1038 <TAK>
A;Cross-references: UNIPROT:P13612; GB:X16983; EMBL:X15356; NID:g33945; PIDN:CAA34852.1
R;Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A;Title: Characterization of the alpha-4 integrin gene promoter.
A;Reference number: A39355; MUID:91239513; PMID:2034655
A;Accession: A39355
A;Molecule type: DNA
A;Residues: 1-93 <ROS>
A;Cross-references: GB:M62841
R;Takada, Y.; Strominger, J.L.; Hemler, M.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987
A;Title: The very late antigen family of heterodimers is part of a superfamily of molec
A;Reference number: A94151; MUID:87204112; PMID:3033641
A;Accession: D28018
A;Molecule type: protein
A;Residues: 40-50,'E',52-53 <TA2>
C;Genetics:
A;Gene: GDB:ITGA4; CD49D
A;Cross-references: GDB:128032; OMIM:192975
A;Map position: 2q31-2q32
C;Superfamily: integrin alpha-4 chain
C;Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transm
F;1-39/Domain: signal sequence #status predicted <SIG>
F;40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

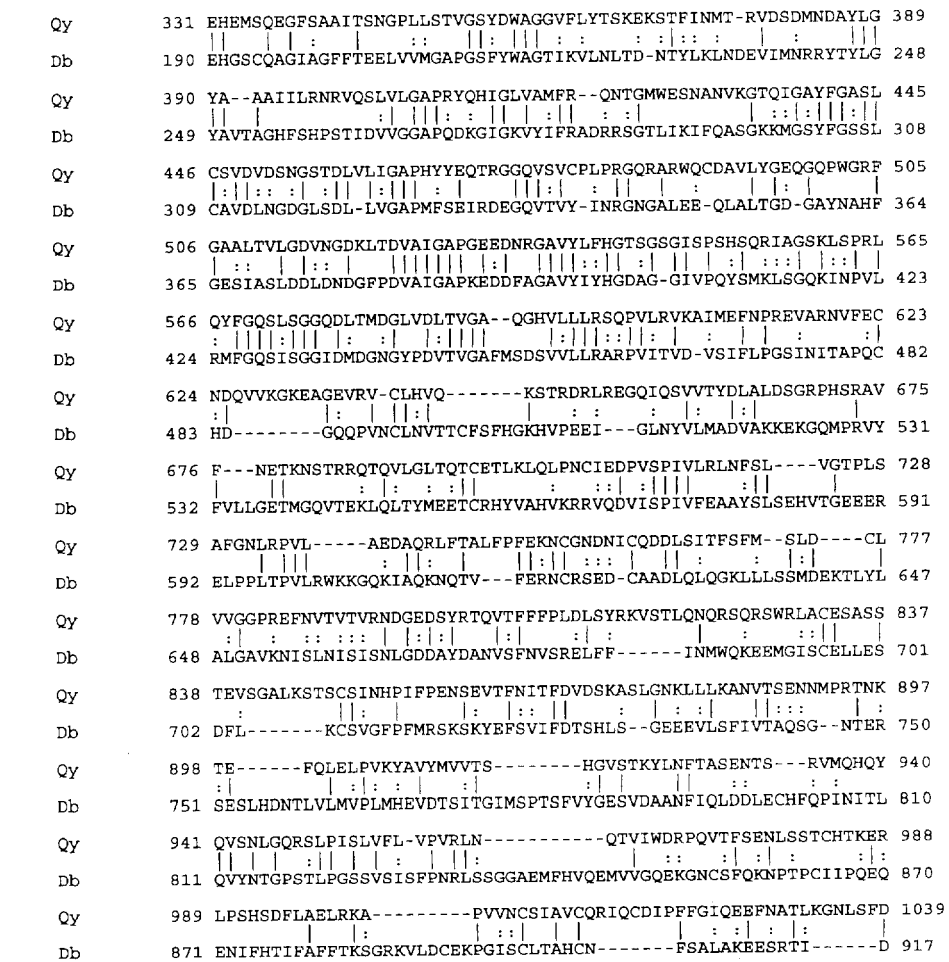
Query Match 10.9%; Score 642; DB 2; Length 1038;
Best Local Similarity 23.8%; Pred. No. 3.8e-36;
Matches 293; Conservative 175; Mismatches 409; Indels 356; Gaps 53;



RESULT 14
I58409
integrin alpha-9 chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I58409; A49459
R;Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah
Oncogene 9, 611-619, 1994
A;Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small
A;Reference number: I58409; MUID:94119603; PMID:8290272
A;Accession: I58409

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1035 <RES>
A;Cross-references: UNIPROT:Q13797; GB:D25303; NID:g464180; PIDN:BAA04984.1; PID:g53332
R;Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.
J. Cell Biol. 123, 1289-1297, 1993
A;Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partn
A;Reference number: A49459; MUID:94064789; PMID:8245132
A;Accession: A49459
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 30-1035 <PAL>
A;Cross-references: GB:L24158
C;Superfamily: integrin alpha-4 chain
C;Keywords: glycoprotein; metal binding; transmembrane protein
F;1-27/Domain: signal sequence #status predicted <SIG>

Query Match 10.7%; Score 630; DB 2; Length 1035;
Best Local Similarity 26.5%; Pred. No. 2.6e-35;
Matches 237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;



```
Qy       1040 WYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEP---------------- 1084
Db        918 IY--------MLLNT-EILKKDSSSVIQ------FMSRAKVKVDPALRVVEIAHGNPEEV 962

Qy       1085 ---FEVPN---PLPLIVG-----SSVGGLLLLALITAALYKLGFFKRQYKDMM 1126
Db        963 TVVFEALHNLEPRGYVVGWIIAISLLVGILIFLLLAVLLWKMGFFRRRYKEII 1015


RESULT 15
T31437
integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C;Species: Lytechinus variegatus (variegated urchin)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31437
R;Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A;Description: Alpha SU2, a sea urchin integrin which binds laminin.
A;Reference number: Z21035
A;Accession: T31437
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1041 <HER>
A;Cross-references: UNIPROT:O76378; EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC2
A;Experimental source: developmental stage embryo
C;Function:
A;Description: binds laminin
C;Superfamily: integrin alpha-2b chain

  Query Match           10.4%;  Score 614.5;  DB 2;  Length 1041;
  Best Local Similarity  24.8%;  Pred. No. 3.1e-34;
  Matches  236;  Conservative  155;  Mismatches  314;  Indels  245;  Gaps   41;

Qy        336 QEGFSAAITSNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDS----DMNDAYLG 389
Db        181 QAGFSGIIFSDNSALVMGAPGSYYLQGQIYVQSLLNRSV-VQATQESNTGTYSFDNSYRG 239

Qy        390 YAAAI--ILRNRVQSLVLGAPRYQHI-GLVAMFRQNTGMWESNANVKGTQIGAYFGASLC 446
Db        240 YSLALGDFNGDGVQDYVVGTPRAESLMGLVAIFDQNLNQFN---QVMGTQIVAYFGYSVT 296

Qy        447 SVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPLPRGQRARWQCDAVLYGEQ-------- 498
Db        297 VVDI-NNDTYDDLLVGAPMYMDDGPAIQ-----------RWEAGAVYVYLQNPDVGPGA 343

Qy        499 -----------GQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGS 546
Db        344 SNRLSLSSTLIGGQIRSRFGLSIASIGDSNQDGFNDVAIGAPYEGDDAGAVYIYHG-SAN 402

Qy        547 GISPSHSQRIAGSKLS-PRLQYFGQSLSGGQDLTMDGLVDLTVGAQ--GHVLLLRSQPVL 603
Db        403 GLKSTPAQVLTPSTLGHSGITTFGFSLQGGQDMDKNKYPDLLVGAESANTAVLIRTRPVV 462

Qy        604 RVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSV----- 658
Db        463 SLDATLNTEP-----------------------IGINLENKTYE-LADGTMVTSFIAMT 497

Qy        659 --------------VTYDLALDSG-RPHSRAVFNETKNSTRRQTQVLGL-TQTCETLKLQ 702
Db        498 CFTYTGNYLPDHIDISYTVTVDSGIIANRRAMFVDNDMSEITKTRRLAVSTQFCDPLRAY 557

Qy        703 LPNCIEDPVSPIVLRLNFSLVGTPLSAFGN---------LRPVLAEDAQRLFTALFPFEK 753
Db        558 VGNSIEDKLTPIKVTLQYDL--------NNDESRLQPHEILPIIDMATMSTQTKQVSIQN 609

Qy        754 NCGNDNICQDDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSY 813
Db        610 NCVN-NICIPDLDVTVT-PNLPNIVIGQTQELTLDVSLNNRGEDAFQSSLSVYYPLGLQF 667

Qy        814 RKVSTLQNQRSQRSWRLACESASSTEVSGALKSTSCSINHPIFPEN------SEVTFNIT 867
Db        668 VRL----ERKANMDFSVTCSEDSD------LRIITCDTGNPMVGKNILEFGLTLSTFQVS 717

Qy        868 FDVDS-----KASLGNKLLLKANVTSENNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTK 922
Db        718 GDKDSIEFYFKAESENS--EDPNTLENNELNMTVPVTVDCTLKLLSASYPEIVMYSTQED 775

Qy        923 YL-------NFTASENTSRVMQHQYQV-----SNLGQRSLPIS----------LVFLVPV 960
Db        776 YVVPPFPAKNASEADIGMEVM-HLYEVRNTGSSNAGEVSLNIQWPQKNEDGEYLFYLLGI 834

Qy        961 RLNQTVIWDRPQ--------------VTFSENLSSTCHTKERLPSHSDFLAELRKAPVV 1005
Db        835 MTEEGVTCQLTQGKANPEGVKLEPSTKAKLSNSTTQVSGRKRREPEVAEALAQTDN--VI 892

Qy       1006 NCSIAVCQRIQCDIPFFGIQEEFNAT------LKGNLSFDWYIKTSHNHLLIVSTAEILF 1059
Db        893 YCASDSCVLINCTI------DEINASKSKVVRILGRF---W--------------ERTF 928

Qy       1060 NDSVFTLLPGQGAFVRSQTETKVE--PFEVPNP----------------LP------- 1092
Db        929 QKAVSELTPVVQATIASSASAAVKTIPYNIPLPRDFSDSTKASTLVTTEELVPPVTPIAW 988

Qy       1093 -LIVGSSVGGLLLLALITAALYKLGFFKR-------QYKDMMS--EGGPP 1132
Db        989 WIIVVSVLGGIILLLIIILGLWKCGFFERKKPGEEKEYAPVASADKDGPP 1038


Search completed: November  9, 2004, 12:19:28
Job time : 33.25 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:46:42 ; Search time 136.75 Seconds
(without alignments)
4783.919 Million cell updates/sec

Title: US-09-902-481B-4
Perfect score: 5884
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5836.5	99.2	1152	1	ITAM_HUMAN	P11215 homo sapien
2	4460	75.8	1153	1	ITAM_MOUSE	P05555 mus musculu
3	4371	74.3	1151	2	Q9JI30	Q9ji30 rattus norv
4	3799.5	64.6	920	2	Q28984	Q28984 sus scrofa
5	3469	59.0	1163	1	ITAX_HUMAN	P20702 homo sapien
6	3401	57.8	1162	1	ITAD_HUMAN	Q13349 homo sapien
7	3319.5	56.4	1188	2	Q6KAS4	Q6kas4 mus musculu
8	3319.5	56.4	1188	2	BAD21383	Bad21383 mus muscu
9	3310.5	56.3	1169	1	ITAX_MOUSE	Q9qxh4 mus musculu
10	3236.5	55.0	1161	1	ITAD_RAT	Q9qye7 rattus norv
11	1572	26.7	1165	1	ITAL_BOVIN	P61625 bos taurus
12	1547.5	26.3	1170	1	ITAL_HUMAN	P20701 homo sapien
13	1546	26.3	1166	2	Q6TYB8	Q6tyb8 bos taurus
14	1546	26.3	1166	2	AAQ90015	Aaq90015 bos tauru
15	1527.5	26.0	1163	1	ITAL_MOUSE	P24063 mus musculu
16	1523.5	25.9	1161	2	Q9WTV4	Q9wtv4 mus musculu
17	1513	25.7	1160	2	Q9R200	Q9r200 mus musculu
18	1395	23.7	1196	2	Q98TF1	Q98tf1 cyprinus ca
19	1350.5	23.0	1086	2	Q96HB1	Q96hb1 homo sapien
20	1344.5	22.9	1187	2	Q98TF0	Q98tf0 cyprinus ca
21	1269	21.6	927	2	Q8HZV0	Q8hzv0 bos taurus
22	1154.5	19.6	1167	2	O88340	O88340 rattus norv
23	1148.5	19.5	1167	1	ITAE_MOUSE	Q60677 mus musculu
24	1140	19.4	1179	1	ITAE_HUMAN	P38570 homo sapien
25	1109	18.8	1167	2	O88341	O88341 rattus norv
26	1093.5	18.6	1151	1	ITA1_HUMAN	P56199 homo sapien
27	1084.5	18.4	1189	1	ITAH_HUMAN	Q9ukx5 homo sapien
28	1071	18.2	1170	1	ITA2_BOVIN	P53710 bos taurus
29	1057	18.0	1178	1	ITA2_MOUSE	Q62469 mus musculu
30	1057	18.0	1178	2	Q6P1C7	Q6p1c7 mus musculu
31	1057	18.0	1178	2	AAH65139	Aah65139 mus muscu
32	1055	17.9	1180	1	ITA1_RAT	P18614 rattus norv
33	1054	17.9	1181	1	ITA2_HUMAN	P17301 homo sapien
34	1051.5	17.9	1167	1	ITAG_HUMAN	O75578 homo sapien
35	1043.5	17.7	1188	1	ITAH_MOUSE	P61622 mus musculu
36	1043.5	17.7	1188	2	Q7TQC3	Q7tqc3 mus musculu
37	1041	17.7	1171	2	O42094	O42094 gallus gall
38	1032	17.5	1038	2	Q8BS01	Q8bs01 mus musculu
39	1005	17.1	895	2	Q9WUF8	Q9wuf8 mus sp. itg
40	998.5	17.0	1160	2	Q8MKF4	Q8mkf4 felis silve
41	865	14.7	348	2	Q8TES5	Q8tes5 homo sapien
42	860	14.6	1332	2	Q9BPQ8	Q9bpq8 halocynthia
43	796	13.5	205	2	Q63001	Q63001 rattus norv
44	770.5	13.1	304	2	Q6PG66	Q6pg66 mus musculu
45	770.5	13.1	304	2	AAH57200	Aah57200 mus muscu

ALIGNMENTS

RESULT 1
ITAM_HUMAN
ID ITAM_HUMAN STANDARD; PRT; 1152 AA.
AC P11215;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha
DE subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)
DE (Neutrophil adherence receptor).
GN Name=ITGAM; Synonyms=CR3A, CD11B;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88315033; PubMed=2457584;
RA Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;
RT "The human leukocyte adhesion glycoprotein Mac-1 (complement receptor
RT type 3, CD11b) alpha subunit. Cloning, primary structure, and relation
RT to the integrins, von Willebrand factor and factor B.";
RL J. Biol. Chem. 263:12403-12411(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88190151; PubMed=2833753;
RA Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;
RT "Molecular cloning of the alpha subunit of human and guinea pig
RT leukocyte adhesion glycoprotein Mo1: chromosomal localization and
RT homology to the alpha subunits of integrins.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88257215; PubMed=2454931;
RA Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;
RT "Amino acid sequence of the alpha subunit of human leukocyte adhesion
RT receptor Mo1 (complement receptor type 3).";
RL J. Cell Biol. 106:2153-2158(1988).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93123748; PubMed=8419480;
RA Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;
RT "Structural analysis of the CD11b gene and phylogenetic analysis of
RT the alpha-integrin gene family demonstrate remarkable conservation of
RT genomic organization and suggest early diversification during
RT evolution.";
RL J. Immunol. 150:480-490(1993).
RN [5]
RP SEQUENCE OF 9-1153 FROM N.A.
RX MEDLINE=89098893; PubMed=2563162;
RA Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L.,
RA Roth G.J.;
RT "cDNA sequence for the alpha M subunit of the human neutrophil
RT adherence receptor indicates homology to integrin alpha subunits.";

```
RL   Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).
RN   [6]
RP   SEQUENCE OF 1-9 FROM N.A.
RX   MEDLINE=92073318; PubMed=1683702;
RA   Shelley C.S., Arnaout M.A.;
RT   "The promoter of the CD11b gene directs myeloid-specific and
RT   developmentally regulated expression.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN   [7]
RP   SEQUENCE OF 1-9 FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=92144986; PubMed=1346576;
RA   Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT   "Characterization of the myeloid-specific CD11b promoter.";
RL   Blood 79:865-870(1992).
RN   [8]
RP   SEQUENCE OF 17-31.
RX   MEDLINE=87076671; PubMed=3539202;
RA   Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
RT   "N-terminal sequence of human leukocyte glycoprotein Mo1: conservation
RT   across species and homology to platelet IIb/IIIa.";
RL   Biochim. Biophys. Acta 874:368-371(1986).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX   MEDLINE=95171458; PubMed=7867070;
RA   Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT   "Crystal structure of the A domain from the alpha subunit of integrin
RT   CR3 (CD11b/CD18).";
RL   Cell 80:631-638(1995).
RN   [10]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX   MEDLINE=96363671; PubMed=8747460;
RA   Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
RT   "Two conformations of the integrin A-domain (I-domain): a pathway for
RT   activation?";
RL   Structure 3:1333-1340(1995).
RN   [11]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX   MEDLINE=98362595; PubMed=9687375;
RA   Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA   Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
RA   Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
RA   Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT   "Cation binding to the integrin CD11b I domain and activation model
RT   assessment.";
RL   Structure 6:923-935(1998).
RN   [12]
RP   3D-STRUCTURE MODELING OF 17-616.
RX   MEDLINE=98226734; PubMed=9560195;
RA   Oxvig C., Springer T.A.;
RT   "Experimental support for a beta-propeller domain in integrin alpha-
RT   subunits and a calcium binding site on its lower surface.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC   -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various
CC       adhesive interactions of monocytes, macrophages and granulocytes
CC       as well as in mediating the uptake of complement-coated particles.
CC       It is identical with CR-3, the receptor for the iC3b fragment of
CC       the third complement component. It probably recognizes the R-G-D
CC       peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for
CC       fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides
CC       of fibrinogen gamma chain.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-M
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC       granulocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
```


```
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```


```
DR   EMBL; J03925; AAA59544.1; -.
DR   EMBL; M18044; AAA59491.1; -.
DR   EMBL; J04145; AAA59903.1; -.
DR   EMBL; S52227; AAB24821.1; -.
DR   EMBL; S52152; AAB24821.1; JOINED.
DR   EMBL; S52153; AAB24821.1; JOINED.
DR   EMBL; S52154; AAB24821.1; JOINED.
DR   EMBL; S52155; AAB24821.1; JOINED.
DR   EMBL; S52157; AAB24821.1; JOINED.
DR   EMBL; S52159; AAB24821.1; JOINED.
DR   EMBL; S52161; AAB24821.1; JOINED.
DR   EMBL; S52164; AAB24821.1; JOINED.
DR   EMBL; S52165; AAB24821.1; JOINED.
DR   EMBL; S52167; AAB24821.1; JOINED.
DR   EMBL; S52169; AAB24821.1; JOINED.
DR   EMBL; S52170; AAB24821.1; JOINED.
DR   EMBL; S52173; AAB24821.1; JOINED.
DR   EMBL; S52174; AAB24821.1; JOINED.
DR   EMBL; S52180; AAB24821.1; JOINED.
DR   EMBL; S52181; AAB24821.1; JOINED.
DR   EMBL; S52184; AAB24821.1; JOINED.
DR   EMBL; S52189; AAB24821.1; JOINED.
DR   EMBL; S52191; AAB24821.1; JOINED.
DR   EMBL; S52192; AAB24821.1; JOINED.
DR   EMBL; S52203; AAB24821.1; JOINED.
DR   EMBL; S52212; AAB24821.1; JOINED.
DR   EMBL; S52213; AAB24821.1; JOINED.
DR   EMBL; S52216; AAB24821.1; JOINED.
DR   EMBL; S52219; AAB24821.1; JOINED.
DR   EMBL; S52220; AAB24821.1; JOINED.
DR   EMBL; S52221; AAB24821.1; JOINED.
DR   EMBL; S52222; AAB24821.1; JOINED.
DR   EMBL; S52226; AAB24821.1; JOINED.
DR   EMBL; M76724; AAA58410.1; -.
DR   EMBL; M84477; AAA51960.1; -.
DR   PIR; A31108; RWHU1B.
DR   PDB; 1A8X; Model; @=17-1152.
DR   PDB; 1BHO; X-ray; 1/2=-.
DR   PDB; 1BHQ; X-ray; 1/2=-.
DR   PDB; 1IDN; X-ray; 1/2=-.
DR   PDB; 1IDO; X-ray; @=140-331.
DR   PDB; 1JLM; X-ray; @=143-334.
DR   PDB; 1M1U; X-ray; A=137-331.
DR   PDB; 1MF7; X-ray; A=144-337.
DR   PDB; 1N9Z; X-ray; A=140-335.
DR   PDB; 1NA5; X-ray; A=144-345.
DR   Genew; HGNC:6149; ITGAM.
DR   MIM; 120980; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0007155; P:cell adhesion; TAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
KW   Glycoprotein; Integrin; Magnesium; Receptor; Repeat; Signal;
KW   Transmembrane.
```

```
FT   SIGNAL        1     16
FT   CHAIN        17   1152       Integrin alpha-M.

  Query Match             99.2%;  Score 5836.5;  DB 1;  Length 1152;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 1127;  Conservative    7;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| ||||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 2
ITAM_MOUSE
ID   ITAM_MOUSE     STANDARD;      PRT;  1153 AA.
AC   P05555; Q8CA73;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha
DE   subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).
GN   Name=Itgam;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=88312584; PubMed=3044779;
RA   Pytela R.;
RT   "Amino acid sequence of the murine Mac-1 alpha chain reveals homology
RT   with the integrin family and an additional domain related to von
RT   Willebrand factor.";
RL   EMBO J. 7:1371-1378(1988).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [3]
RP   SEQUENCE OF 11-45 FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Spleen;
RX   MEDLINE=86287312; PubMed=2942940;
RA   Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E.,
RA   Larson R.S., Roberts T.M., Springer T.A.;
RT   "A partial genomic DNA clone for the alpha subunit of the mouse
```

```
RT   complement receptor type 3 and cellular adhesion molecule Mac-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).
RN   [4]
RP   SEQUENCE OF 17-28.
RX   MEDLINE=85188276; PubMed=3887182;
RA   Springer T.A., Teplow D.B., Dreyer W.J.;
RT   "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
RT   glycoproteins and unexpected relation to leukocyte interferon.";
RL   Nature 314:540-542(1985).
CC   -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various
CC       adhesive interactions of monocytes, macrophages and granulocytes
CC       as well as in mediating the uptake of complement-coated particles.
CC       It is identical with CR-3, the receptor for the iC3b fragment of
CC       the third complement component. It probably recognizes the R-G-D
CC       peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for
CC       fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides
CC       of fibrinogen gamma chain. Alpha-M/beta-2 play a critical role in
CC       mast cell development and in immune complex-mediated
CC       glomerulonephritis. Mice expressing a null mutation of the alpha-M
CC       subunit gene demonstrate increase in neutrophil accumulation, in
CC       response to a impaired degranulation and phagocytosis, events that
CC       apparently accelerate apoptosis in neutrophils. These mice develop
CC       obesity.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta chain. Alpha-M
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P05555-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P05555-2; Sequence=VSP_010473;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC       granulocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X07640; CAA30479.1; -.
DR   EMBL; AK039444; BAC30350.1; -.
DR   EMBL; M14293; AAA39484.1; -.
DR   PIR; S00551; S00551.
DR   HSSP; P11215; 1BHQ.
DR   MGD; MGI:96607; Itgam.
DR   GO; GO:0009897; C:external side of plasma membrane; IDA.
DR   GO; GO:0007155; P:cell adhesion; IMP.
DR   GO; GO:0045123; P:cellular extravasation; IMP.
DR   GO; GO:0030593; P:neutrophil chemotaxis; IMP.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Alternative splicing; Calcium; Cell adhesion;
KW   Direct protein sequencing; Glycoprotein; Integrin; Receptor; Repeat;
KW   Signal; Transmembrane.
```

```
FT   SIGNAL       1    16
FT   CHAIN       17  1153       Integrin alpha-M.
FT   DOMAIN      17  1105       Extracellular (Potential).
FT   TRANSMEM  1106  1129       Potential.
FT   DOMAIN    1130  1153       Cytoplasmic (Potential).
FT   REPEAT      31    84       FG-GAP 1.
FT   REPEAT      85   163       FG-GAP 2.
FT   DOMAIN     164   350       VWFA.
FT   REPEAT     337   400       FG-GAP 3.
FT   REPEAT     401   452       FG-GAP 4.
FT   REPEAT     454   515       FG-GAP 5.
FT   REPEAT     517   575       FG-GAP 6.
FT   REPEAT     580   632       FG-GAP 7.
FT   CA_BIND    465   473       Potential.
FT   CA_BIND    529   537       Potential.
FT   CA_BIND    592   600       Potential.
FT   SITE      1132  1136       GFFKR motif.
FT   DISULFID    66    73       By similarity.
FT   DISULFID   105   123       By similarity.
FT   DISULFID   654   711       By similarity.
FT   DISULFID   770   776       By similarity.
FT   DISULFID   999  1023       By similarity.
FT   DISULFID  1028  1033       By similarity.
FT   CARBOHYD    58    58       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    86    86       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   391   391       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   696   696       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   734   734       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   772   772       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   801   801       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   881   881       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   907   907       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   941   941       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   980   980       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   994   994       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1022  1022       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1045  1045       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1051  1051       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1076  1076       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC   453   569       Missing (in isoform 2).
FT                              /FTId=VSP_010473.
FT   CONFLICT    37    37       N -> S (in Ref. 2).
FT   CONFLICT   683   683       V -> G (in Ref. 2).
SQ   SEQUENCE  1153 AA;  127480 MW;  178DB988AECB0343 CRC64;
```

```
  Query Match             75.8%;  Score 4460;  DB 1;  Length 1153;
  Best Local Similarity   73.8%;  Pred. No. 4.1e-288;
  Matches  840;  Conservative  146;  Mismatches  150;  Indels    2;  Gaps    2;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||| |||||||:||||:|||| |: ||| |||| | || |:||||||| | ||
Db         17 FNLDTEHPMTFQENAKGFGQNVVQLGGTSVVVAAPQEAKAVNQTGALYQCDYSTSRCHPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||| :| | |||||||||| | ||||| |||:|||||| : ||:
Db         77 PLQVPPEAVNMSLGLSLAVSTVPQQLLACGPTVHQNCKENTYVNGLCYLFGSNLLRPPQQ 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              |||||| |||::||| |||||||||    ||::|||:||||||| ||||||||||||:||
Db        137 FPEALRECPQQESDIVFLIDGSGSINNIDFQKMKEFVSTVMEQFKKSKTLFSLMQYSDEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||| :|: ||:||| : || || ||| ||:|:|||||||: |||||:|| ||| ::
Db        197 RIHFTFNDFKRNPSPRSHVSPIKQLNGRTKTASGIRKVVRELFHKTNGARENAAKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||| |:||||| || |||||||||:||    :||:||:|:|||| :||||::
Db        257 TDGEKFGDPLDYKDVIPEADRAGVIRYVIGVGNAFNKPQSRRELDTIASKPAGEHVFQVD 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||| ||||||:|||||||||||||:|||||||||||||||:|||| :|||:||||
Db        317 NFEALNTIQNQLQEKIFAIEGTQTGSTSSFEHEMSQEGFSASITSNGPLLGSVGSFDWAG 376
```

```
Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        377 GAFLYTSKDKVTFINTTRVDSDMNDAYLGYASAVILRNRVQSLVLGAPRYQHIGLVVMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        437 ENFGTWEPHTSIKGSQIGSYFGASLCSVDMDADGNTNLILIGAPHYYEKTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        497 PRG-RARWQCEALLHGDQGHPWGRFGAALTVLGDVNGDKLTDVAIGAPGEQENQGAVYIF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        556 YGASIASLSASHSHRIIGAHFSPGLQYFGQSLSGGKDLTMDGLMDLAVGAQGHLLLLRAQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        616 PVLRLEATMEFSPKKVARSVFACQEQVLKNKDAGEVRVCLRVRKNTKDRLREGDIQSTVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        676 YDLALDPVRSRIRAFFDETKNNTRRRTQVFGLMQKCETLKLILPDCVDDSVSPIILRLNY 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        736 TLVGEPLRSFGNLRPVLAMDAQRFFTAMFPFEKNCGNDSICQDDLSITMSAMGLDTLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRL-ACESASSTE 839
Db        796 GPQDFNMSVTLRNDGEDSYGTQVTVYYPSGLSYRKDSASQNPLTKKPWFVKPAESSSSSE 855

Qy        840 VSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTE 899
Db        856 GHGALKSTTWNININHPIFPANSEVTFNVTFDVDSHASFGNKLLLKAIVASENNMSRTHKTK 915

Qy        900 FQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        916 FQLELPVKYAIYMIVTSDESSIRYLNFTASEMTSKVIQHQYQFNNLGQRSLPVSVVFWIP 975

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        976 VQINNVTVWDHPQVIFSQNLSSACHTEQKSPPHSNFRDQLERTPVLNCSVAVCKRIQCDL 1035

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1036 PSFNTQEIFNVTLKGNLSFDWYIKTSHGHLLLVSSTEILFNDSAFALLPGQESYVRSKTE 1095

Qy       1080 TKVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1096 TKVEPYEVHNPVPLIVGSSIGGLVLLALITAGLYKLGFFKRQYKDMMNEAAPQDAPPQ 1153
```

```
RESULT 3
Q9JI30
ID   Q9JI30      PRELIMINARY;      PRT;  1151 AA.
AC   Q9JI30;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Integrin beta 2 alpha subunit.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fathallah D.M. Sr., Zerria K. Jr.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR   EMBL; AF268593; AAF81280.1; -.
DR   HSSP; P11215; 1BHQ.
DR   GO; GO:0008305; C:integrin complex; IEA.
DR   GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR   GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 2.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Cell adhesion; Integrin; Transmembrane.
SQ   SEQUENCE   1151 AA;  126943 MW;  8F785695D4074CA5 CRC64;

  Query Match           74.3%;  Score 4371;  DB 2;  Length 1151;
  Best Local Similarity  72.6%;  Pred. No. 3.5e-282;
  Matches  826;  Conservative  150;  Mismatches  159;  Indels    2;  Gaps    2;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         17 FNLDTENPMTFQENASGFGQSVIQLGETRVVVAAPQEVKAVNQTGALYQCDYSTNRCDPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         77 PLQVPPEAVNMSLGLSLAATTVPPQLLACGPTVHQNCKENTYVNGLCYLFGSNLLRKPQQ 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        137 FPEALRGCPQQESNIAFLIDGSGSINTIDFQKMKEFVSTVMDQFQKSKTLFSLMQYSDEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db        197 RTHFTFNDFKRNPDPKSHVRPIRQLNGRTKTASGIRKVVRELFQKINGARDNAAKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db        257 TDGEKFGDPLNYEDVIPEAEEAGIIRYVIGVGNAFHKPQSRRELDTIASKPAGDHVFQVD 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        317 NFEALNTIRNQLQEKIFAIEGTQTGSTSSFEHEMSQEGFSAAITSNGPLLGSVGSFDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        377 GAFLYPSKDKASFINTTRIDSDMNDAYLGYASAVISRNRVQSLVLGAPRYQHIGLVVMFK 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        437 QNFGAWEPHTDIKGSQIGSYFGASLCSVDMDADGNTNLILIGAPHHYEKTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        497 PRG-RARWQCEAILHGDQGHPWGRFGAALTVLGDVNGDKLTDVAIGAPGEQENQGAVYIF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        556 HGASVASISTPHSQRIAGARFSPGLQYFGQSLSGGKDLTMDGLMDLAVGAQGRLLLIRAQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        616 PVLRLEATMEFSPKNLARSVFACQEEVVKNKDAGEVRVCLQVRKNTKDRLREGDIQSTVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        676 YDLALDPGRSVVRAFFDETKNGILRRIRVFGLTQKCETLKLILPDCVDNSVSPIILRLNY 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        736 TLVGEPLRSSRDLRPVLAMEAQRIFTAMFPFEKNCGNDTICQDDLSITVSSTRVNTLVVG 795
```

```
Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        796 DSRDFDVSVTLRNDGEDSYGTKVTCYYPSGLSYRKVSASQNQFSKKPWRVIAE-PSSSEG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 QGVLKSTIWDINHPIFPANSEVTFNVTFDVDSDASLGNKLLLKVVVTSENNVSRTDKTES 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        915 QLELPVKYAIYMVVTSGESSIKYLNFTASEMTSKVIQHQYQFNNLGQRSFPVSVVFWIPV 974

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        975 QINKVTIWDPPQVTFSQNLSSVCRTEQKSPSHSKFQDELERTPVLNCSVAVCKRIQCDIP 1034

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1035 SFNSKEIFNVTLQGNLLFDWYIETSHDHLLLVSTAEILFNDSAFALLPGQETFVKAQTET 1094

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1095 KVEPYTVHNPVPLIVGSSVGGLVLLALITAGLYKLGFFKRQYKDMMNEAGGQDGPPQ 1151
```

RESULT 4
Q28984
ID Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CD11b (Fragment).
GN Name=CD11b;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; U40072; AAB16869.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Cell adhesion; Integrin; Transmembrane.
FT NON_TER 1 1
FT NON_TER 920 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;

Query Match 64.6%; Score 3799.5; DB 2; Length 920;
Best Local Similarity 78.8%; Pred. No. 3.2e-244;
Matches 726; Conservative 84; Mismatches 110; Indels 1; Gaps 1;

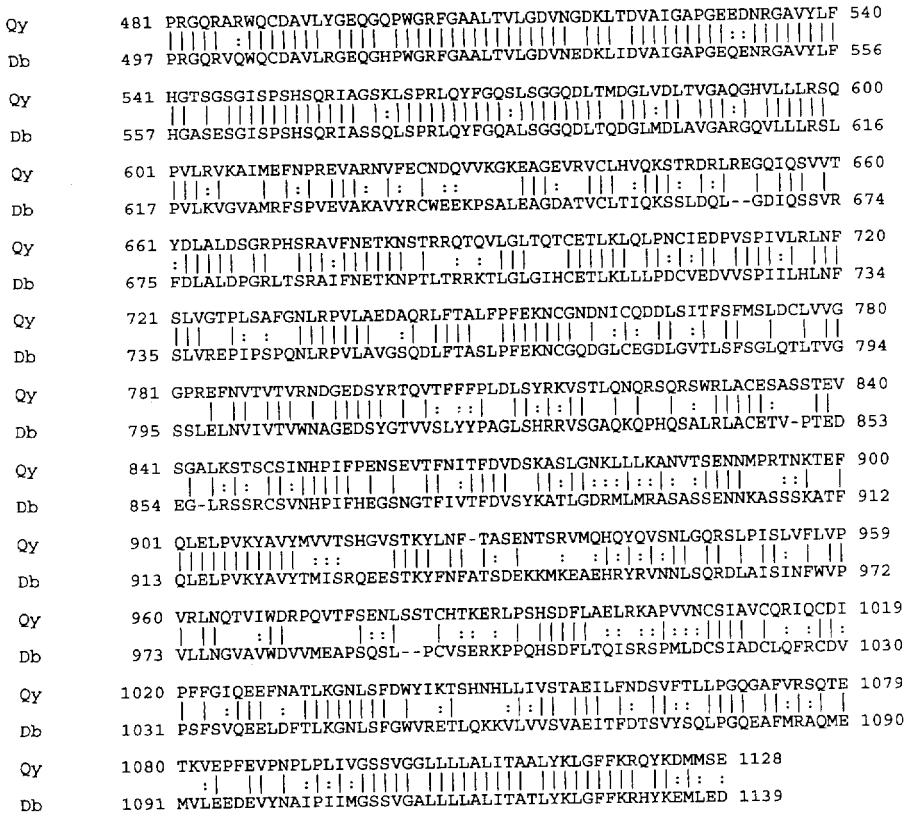
```
Qy        118 PQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYS 177
Db          1 PQKFPEALRGCPQQESDIAFLIDGSGSINRLDFQRMKEFVSTVMGQFQKSKTLFALMQYS 60

Qy        178 EEFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKIL 237
Db         61 EDFYTHFTFNDFKRNPSPKLLVRPIRQLLGRTHTATGIRKVVRELFHSKSGARENALKIL 120

Qy        238 FLLTDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVF 297
Db        121 VVITDGEKFGDPLGYEDVIPEADRKGVIRYVIGVGDAFNSWKSREELNTIASKPCGDHVF 180

Qy        298 QINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYD 357
Db        181 QVNNFEAVKTIQNQLQEKTFAIEGTQTGSTSSFECEMSQEGFSAAITSNGPLLGAVGSFD 240

Qy        358 WAGGVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVA 417
Db        241 WAGGAFLHMPKDRVIFINTTRVDSDMNDAYLGYAVEVILRNQAQSLVLGAPRYQHTGLVV 300

Qy        418 MFRQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSV 477
Db        301 MFKQNSGAWEKNADIKGSQIGSYFGASLCSVDVNRDGSSDLVLIGAPHYYEQTRGGQVSV 360

Qy        478 CPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAV 537
Db        361 CPLPQG-RAKWQCRVILCGEQGHPWSRFGAALTALGDVNGDKLTDVAIGAPGEQDNRGAV 419

Qy        538 YLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLL 597
Db        420 YLFHGTSELGISPSHSQRIAGSQLSPRLQYFGQSLSGGQDLTMDGLMDLAVGAQGHVFLL 479

Qy        598 RSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQS 657
Db        480 RSQPVLRVEASMVFKPREVARNVYECRQQAAKTQIAGEVQVCLQVRKSTWDRLREGDTQS 539

Qy        658 VVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLR 717
Db        540 IITYDLALDPGRPHPRAVFEETKNNTRRQTQTLGLSRKCEHLALWLPDCVEDSVTPIVLR 599

Qy        718 LNFSLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCL 777
Db        600 LNFSLVGKPASSFGNLRPVLAVDAQRLFTALFPFEKNCGNDSICQDDLSITFSFMSLDTL 659

Qy        778 VVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASS 837
Db        660 VVGGPRDLKVTLTVRNQGEDSYRTQVTFFYPSDLSYRKVSTSQNQRSQRSWRLACESDVS 719

Qy        838 TEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNK 897
Db        720 TEESTALKSTSCSINHPIFPDNSEVTFNVTFDVNPDAFLGYKLLLKANVTSENNMPSSNK 779

Qy        898 TEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFL 957
Db        780 TEFQLELPVKYAVYVVVTSLEVSTKYFNFTASEKTRHVIEHQYQFNNLGQRKLPISVVFW 839

Qy        958 VPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQC 1017
Db        840 VPVRLNRVTVWDQPQVTFSQNLSRSCSTEEIGPRHSDFLEKLQKTPVLNCSIAVCQKIQC 899

Qy       1018 DIPFFGIQEEFNATLKGNLSF 1038
Db        900 DIPSFGIQEELKVTLKGNLSF 920
```

RESULT 5
ITAX_HUMAN
ID ITAX_HUMAN STANDARD; PRT; 1163 AA.
AC P20702; Q8IVA6;
DT 01-FEB-1991 (Rel. 17, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
GN Name=ITGAX; Synonyms=CD11C;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

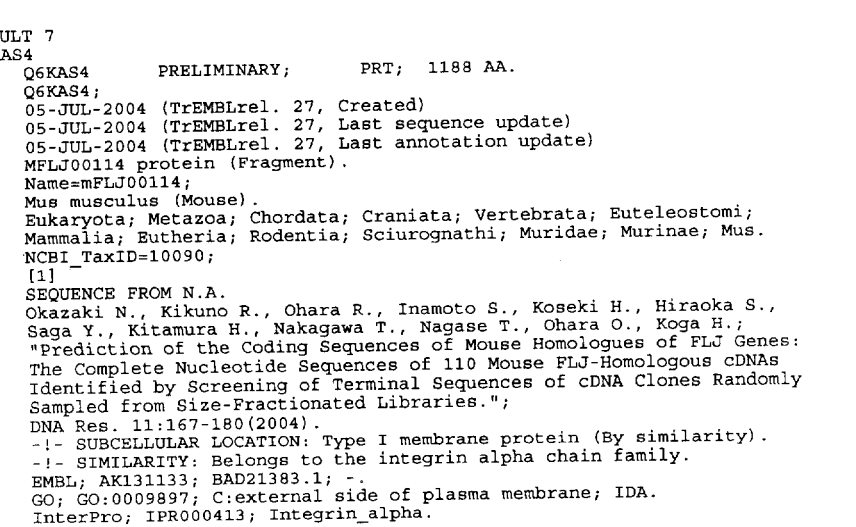
```
RX   MEDLINE=88166645; PubMed=3327687;
RA   Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT   "cDNA cloning and complete primary structure of the alpha subunit of a
RT   leukocyte adhesion glycoprotein, p150,95.";
RL   EMBO J. 6:4023-4028(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90153906; PubMed=2303426;
RA   Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT   "Genomic structure of an integrin alpha subunit, the leukocyte p150,95
RT   molecule.";
RL   J. Biol. Chem. 265:2782-2788(1990).
RN   [3]
RP   ERRATUM.
RA   Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL   J. Biol. Chem. 265:12750-12751(1990).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 20-43.
RX   MEDLINE=87167596; PubMed=3549901;
RA   Miller L.J., Wiebe M., Springer T.A.;
RT   "Purification and alpha subunit N-terminal sequences of human Mac-1
RT   and p150,95 leukocyte adhesion proteins.";
RL   J. Immunol. 138:2381-2383(1987).
CC   -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
CC       recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
CC       interaction during inflammatory responses. It is especially
CC       important in monocyte adhesion and chemotaxis.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC       granulocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M81695; AAA59180.1; -.
DR   EMBL; M29165; -; NOT_ANNOTATED_CDS.
DR   EMBL; M29487; AAA51620.1; ALT_SEQ.
DR   EMBL; M29482; AAA51620.1; JOINED.
DR   EMBL; M29483; AAA51620.1; JOINED.
DR   EMBL; M29484; AAA51620.1; JOINED.
DR   EMBL; M29485; AAA51620.1; JOINED.
DR   EMBL; M29486; AAA51620.1; JOINED.
DR   EMBL; BC038237; AAH38237.1; -.
DR   PIR; A36584; RWHU1C.
DR   PDB; 1N3Y; X-ray; A=141-338.
DR   Genew; HGNC:6152; ITGAX.
DR   MIM; 151510; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0004872; F:receptor activity; TAS.
DR   GO; GO:0007155; P:cell adhesion; TAS.
DR   GO; GO:0009887; P:organogenesis; TAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
KW   Glycoprotein; Integrin; Magnesium; Polymorphism; Receptor; Repeat;
KW   Signal; Transmembrane.
FT   SIGNAL       1     19
FT   CHAIN       20   1163       Integrin alpha-X.
FT   DOMAIN      20   1107       Extracellular (Potential).
FT   TRANSMEM  1108   1128       Potential.
FT   DOMAIN    1129   1163       Cytoplasmic (Potential).
FT   REPEAT      34     87       FG-GAP 1.
FT   REPEAT      88      ?       FG-GAP 2.
FT   DOMAIN     165    351       VWFA.
FT   REPEAT       ?    401       FG-GAP 3.
FT   REPEAT     402    453       FG-GAP 4.
FT   REPEAT     455    517       FG-GAP 5.
FT   REPEAT     518    576       FG-GAP 6.
FT   REPEAT     581    633       FG-GAP 7.
FT   CA_BIND    466    474       Potential.
FT   CA_BIND    530    538       Potential.
FT   CA_BIND    593    601       Potential.
FT   SITE      1131   1135       GFFKR motif.
FT   DISULFID    69     76       By similarity.
FT   DISULFID   108    126       By similarity.
FT   DISULFID   655    712       By similarity.
FT   DISULFID   771    777       By similarity.
FT   DISULFID   848    863       By similarity.
FT   DISULFID   998   1022       By similarity.
FT   DISULFID  1027   1032       By similarity.
FT   CARBOHYD    61     61       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   392    392       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   697    697       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   735    735       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   899    899       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   939    939       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1050   1050       N-linked (GlcNAc. . .) (Potential).
FT   VARIANT     48     48       W -> R (in dbSNP:11574633).
FT                               /FTId=VAR_018672.
FT   CONFLICT   209    209       T -> S (in Ref. 4).
FT   CONFLICT   251    251       T -> A (in Ref. 4).
FT   CONFLICT   469    469       T -> S (in Ref. 4).
FT   CONFLICT   490    490       G -> A (in Ref. 2).
FT   CONFLICT   547    547       E -> K (in Ref. 4).
FT   CONFLICT   756    756       D -> L (in Ref. 1).
FT   CONFLICT   819    819       I -> V (in Ref. 4).
FT   CONFLICT  1161   1163       SEK -> TPHYPQDNV (in Ref. 4).
FT   STRAND     150    157
```

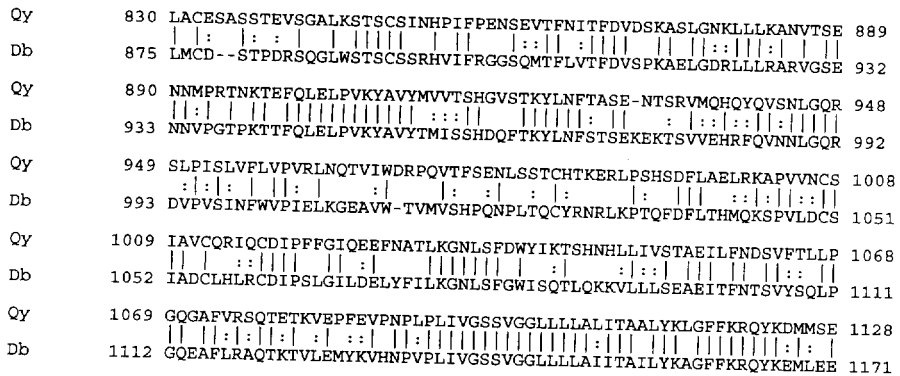
```
FT   TURN        160    161
FT   HELIX       164    178
FT   TURN        179    180
FT   TURN        183    185
FT   STRAND      186    193
FT   STRAND      197    201
FT   HELIX       203    208
FT   HELIX       212    216
FT   TURN        217    218
FT   STRAND      226    226
FT   HELIX       228    236
FT   TURN        237    240
FT   HELIX       242    244
FT   TURN        245    245
FT   TURN        248    249
FT   STRAND      251    258
FT   STRAND      263    263
FT   HELIX       269    278
FT   TURN        279    280
FT   STRAND      282    288
FT   HELIX       290    293
FT   TURN        296    297
FT   HELIX       298    304
FT   HELIX       310    312
FT   STRAND      313    316
FT   HELIX       319    325
FT   HELIX       326    334
FT   TURN        335    335
SQ   SEQUENCE   1163 AA;  127886 MW;  83658A13B5C5DE8F CRC64;

  Query Match             59.0%;  Score 3469;  DB 1;  Length 1163;
  Best Local Similarity   60.9%;  Pred. No. 5.1e-222;
  Matches  688;  Conservative  142;  Mismatches  293;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||     |: :: ||| ||||   | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               |||| ||||||||||||:|||| |||||||||  |  | |: |||||  |    |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
               | : : ||::: || |||||||||   :|  |  :|  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              : ||||:||:   || ||:  : || | |:||| :: ||  ||: : |||::| ||| ::
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:| || | |:|||| |  |:||| |||| ||::  | :||| :|||| ::|:|::
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:|||||||||:| |||||| ||:|||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||       |||||:: : || |:||||:  : |   |||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:|||||||||||::|||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| ||||||||||||||||||||| |||||||:|||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  ||||||||||||:|| |||||||:||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :||| :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
               ||||| ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:|| |||| ||| |||||| |:||||:| |:||| | |:||
Db        737 TLVGKPLLAFRNLRPMLAADAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
                | |  | | ||||||| | :||  |  |||| |:  | |  ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::|| |||| || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||||| ||:||  |||||: | |  | |  |:|||:|||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| |  | ::::  |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :     : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143


RESULT 6
ITAD_HUMAN
ID   ITAD_HUMAN     STANDARD;      PRT;  1162 AA.
AC   Q13349; Q15575; Q15576;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN   Name=ITGAD;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=96111956; PubMed=8777714;
RA   Van der Vieren M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA   Staunton D.E., Gallatin W.M.;
RT   "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
RT   3.";
RL   Immunity 3:683-690(1995).
RN   [2]
RP   SEQUENCE OF 1-235 FROM N.A.
RX   MEDLINE=20187620; PubMed=10722744;
RA   Noti J.D., Johnson A.K., Dillon J.D.;
RT   "Structural and functional characterization of the leukocyte integrin
RT   gene CD11d. Essential role of Sp1 and Sp3.";
RL   J. Biol. Chem. 275:8959-8969(2000).
RN   [3]
RP   SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX   MEDLINE=96257236; PubMed=8666289;
RA   Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
RT   "Cloning and chromosomal localization of a novel gene-encoding a human
RT   beta 2-integrin alpha subunit.";
```

```
RL   Gene 171:291-294(1996).
RN   [4]
RP   INTERACTION WITH VCAM1.
RX   MEDLINE=99059842; PubMed=9841932;
RA   Grayson M.H., Van der Vieren M., Sterbinsky S.A., Michael Gallatin W.,
RA   Hoffman P.A., Staunton D.E., Bochner B.S.;
RT   "alphadbeta2 integrin is expressed on human eosinophils and functions
RT   as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-
RT   1).";
RL   J. Exp. Med. 188:2187-2191(1998).
RN   [5]
RP   INTERACTION WITH VCAM1.
RX   MEDLINE=99370002; PubMed=10438935;
RA   Van der Vieren M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
RA   Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RT   "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
RT   binding interface between I domain and VCAM-1.";
RL   J. Immunol. 163:1984-1990(1999).
CC   -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC       VCAM1. May play a role in the atherosclerotic process such as
CC       clearing lipoproteins from plaques and in phagocytosis of blood-
CC       borne pathogens, particulate matter, and senescent erythrocytes
CC       from the blood.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell
CC       lines and subsets of peripheral blood leukocytes and strongly on
CC       tissue-specialized cells, including macrophages foam cells within
CC       atherosclerotic plaques, and on splenic red pulp macrophages.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U37028; AAB38547.1; -.
DR   EMBL; U40274; AAB60634.1; -.
DR   EMBL; U40275; AAB60635.1; -.
DR   EMBL; U40276; AAB60636.1; -.
DR   EMBL; U40277; AAB60637.1; -.
DR   EMBL; U40279; AAB60638.1; -.
DR   EMBL; U40278; AAB60638.1; JOINED.
DR   EMBL; AF187881; AAF62875.1; -.
DR   HSSP; P11215; 1BHQ.
DR   Genew; HGNC:6146; ITGAD.
DR   MIM; 602453; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0016337; P:cell-cell adhesion; NAS.
DR   GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     17       Potential.
FT   CHAIN        18   1162       Integrin alpha-D.
FT   DOMAIN       18   1100       Extracellular (Potential).
FT   TRANSMEM   1101   1121       Potential.
FT   DOMAIN     1122   1162       Cytoplasmic (Potential).
FT   REPEAT       32     85       FG-GAP 1.
FT   REPEAT       86      ?       FG-GAP 2.
FT   DOMAIN      150    332       VWFA.
FT   REPEAT      350    400       FG-GAP 3.
FT   REPEAT      401    452       FG-GAP 4.
FT   REPEAT      454    516       FG-GAP 5.
FT   REPEAT      518    576       FG-GAP 6.
FT   REPEAT      581    633       FG-GAP 7.
FT   CA_BIND     465    473       Potential.
FT   CA_BIND     530    538       Potential.
FT   CA_BIND     593    601       Potential.
FT   SITE       1127   1131       GFFKR motif.
FT   DISULFID     67     74       By similarity.
FT   DISULFID    106    124       By similarity.
FT   DISULFID    655    710       By similarity.
FT   DISULFID    769    775       By similarity.
FT   DISULFID    846    861       By similarity.
FT   DISULFID    994   1018       By similarity.
FT   DISULFID   1023   1028       By similarity.
FT   CARBOHYD      59     59       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD      87     87       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD      99     99       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     391    391       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     691    691       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     733    733       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     873    873       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     957    957       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    1046   1046       N-linked (GlcNAc. . .) (Potential).
FT   CONFLICT     500    500       Missing (in Ref. 2).
FT   CONFLICT     515    518       GHPW -> ATP (in Ref. 2).
FT   CONFLICT     825    825       L -> V (in Ref. 2).
FT   CONFLICT     984    984       V -> A (in Ref. 2).
SQ   SEQUENCE   1162 AA;  126885 MW;  F296A1A35455D77D CRC64;

  Query Match             57.8%;  Score 3401;  DB 1;  Length 1162;
  Best Local Similarity   59.3%;  Pred. No. 1.8e-217;
  Matches  669;  Conservative  167;  Mismatches  285;  Indels    8;  Gaps    6;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    |||:| |||||||| |||:||||| |:||||| | || |  :|| |:||
Db         18 FNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQTGRLYDCAAATGMCQPI 77

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              | :  |||||||||:|||:|:  :|||||||:|: | ||:| || | | ||   :  |
Db         78 PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-EIIQT 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              |:|   || :: || |||||||||  :|| :|| :|  || | : : |||:|||||
Db        137 VPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFEGTDTLFALMQYSNLL 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              :||||| :|: :|: :||: || || | | ||||:  || :||:  |||||:| ||| ::
Db        197 KIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHHKNGARKSAKKILIVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:|: ||| | |||| :: |:||| |||| ||:   :||||||::| ||:||||:::
Db        257 TDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELNTISSAPPQDHVFKVD 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              || || :|| ||:|||:|:|||| : ||||:|||||||| |:| :|  |  |||: |:|
Db        317 NFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGSFSWSG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:|||:  : |   ||:|||||||||| |   :|
Db        377 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVLGAPRYQHTGKAVIFT 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:|||||||||||||:|||||:|||||||||||||||||||||
Db        437 QVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPHYYEQTRGGQVSVCPL 496
```



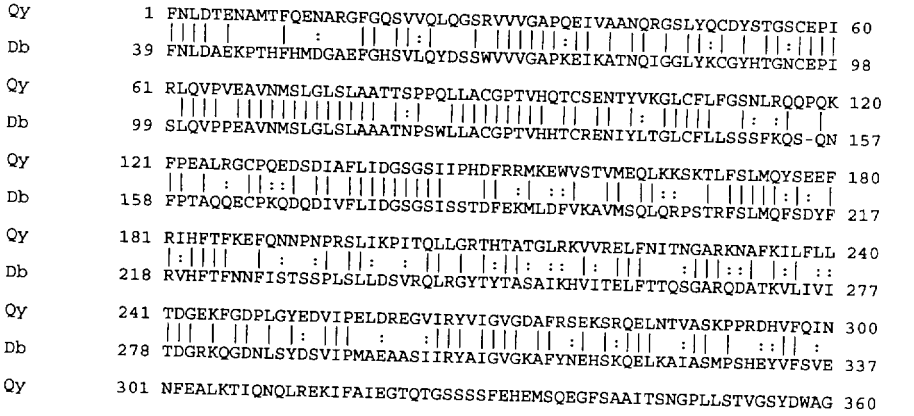
```
RESULT 7
Q6KAS4
ID   Q6KAS4      PRELIMINARY;      PRT;  1188 AA.
AC   Q6KAS4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MFLJ00114 protein (Fragment).
GN   Name=mFLJ00114;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA   Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT   "Prediction of the Coding Sequences of Mouse Homologues of FLJ Genes:
RT   The Complete Nucleotide Sequences of 110 Mouse FLJ-Homologous cDNAs
RT   Identified by Screening of Terminal Sequences of cDNA Clones Randomly
RT   Sampled from Size-Fractionated Libraries.";
RL   DNA Res. 11:167-180(2004).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR   EMBL; AK131133; BAD21383.1; -.
DR   GO; GO:0009897; C:external side of plasma membrane; IDA.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; UNKNOWN_1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Cell adhesion; Integrin; Transmembrane.
FT   NON_TER       1       1
SQ   SEQUENCE   1188 AA;  131248 MW;  B8D93107BDBB4178 CRC64;

  Query Match             56.4%;  Score 3319.5;  DB 2;  Length 1188;
  Best Local Similarity   56.8%;  Pred. No. 4.9e-212;
  Matches  647;  Conservative  171;  Mismatches  303;  Indels   19;  Gaps    7;
```

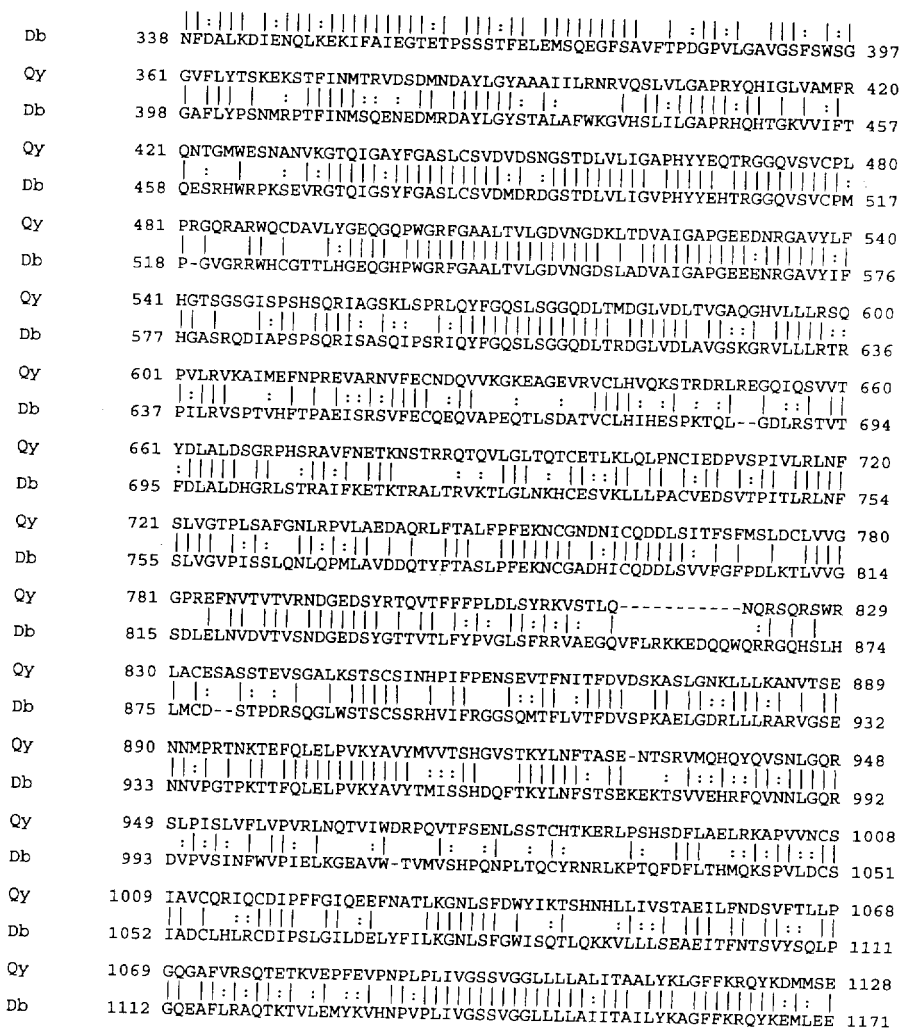





RESULT 8
BAD21383
ID BAD21383 PRELIMINARY; PRT; 1188 AA.
AC BAD21383;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE MFLJ00114 protein (Fragment).
GN MFLJ00114.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Gene:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-homologous cDNAs
RT Identified by Screening of Terminal sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK131133; BAD21383.1; -.
FT NON_TER 1 1
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDBB4178 CRC64;

Query Match 56.4%; Score 3319.5; DB 2; Length 1188;
Best Local Similarity 56.8%; Pred. No. 4.9e-212;
Matches 647; Conservative 171; Mismatches 303; Indels 19; Gaps 7;





RESULT 9
ITAX_MOUSE
ID ITAX_MOUSE STANDARD; PRT; 1169 AA.
AC Q9QXH4;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c).
GN Name=Itgax;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Dendritic cell;

```
RA   Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RA   Tsuchiya H.;
RT   "Isolation of genes selectively expressed by dendritic cells.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
CC       recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
CC       interaction during inflammatory responses. It is especially
CC       important in monocyte adhesion and chemotaxis (By similarity).
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
CC       associates with beta-2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF211864; AAF23492.1; -.
DR   HSSP; P20702; 1N3Y.
DR   MGD; MGI:96609; Itgax.
DR   GO; GO:0009897; C:external side of plasma membrane; IDA.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL       1   19       Potential.
FT   CHAIN       20  1169       Integrin alpha-X.
FT   DOMAIN      20  1116       Extracellular (Potential).
FT   TRANSMEM  1117  1137       Potential.
FT   DOMAIN    1138  1169       Cytoplasmic (Potential).
FT   REPEAT      34    87       FG-GAP 1.
FT   REPEAT      88     ?       FG-GAP 2.
FT   DOMAIN     152   330       VWFA.
FT   REPEAT       ?   402       FG-GAP 3.
FT   REPEAT     403   454       FG-GAP 4.
FT   REPEAT     456   518       FG-GAP 5.
FT   REPEAT     519   577       FG-GAP 6.
FT   REPEAT     582   634       FG-GAP 7.
FT   CA_BIND    467   475       Potential.
FT   CA_BIND    531   539       Potential.
FT   CA_BIND    594   602       Potential.
FT   SITE      1140  1144       GFFKR motif.
FT   DISULFID    69    76       By similarity.
FT   DISULFID   108   126       By similarity.
FT   DISULFID   656   711       By similarity.
FT   DISULFID   770   776       By similarity.
FT   DISULFID   858   873       By similarity.
FT   DISULFID  1007  1031       By similarity.
FT   DISULFID  1036  1041       By similarity.
FT   CARBOHYD    89    89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   267   267       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   393   393       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   734   734       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   949   949       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1059  1059       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1084  1084       N-linked (GlcNAc. . .) (Potential).
```

```
SQ   SEQUENCE   1169 AA;  129150 MW;  C616412033C219A6 CRC64;

  Query Match           56.3%;  Score 3310.5;  DB 1;  Length 1169;
  Best Local Similarity  56.6%;  Pred. No. 1.9e-211;
  Matches  645;  Conservative  173;  Mismatches  303;  Indels   19;  Gaps    7;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    |  :   || ||:|   | ||||||:|| | || | ||:| | ||:|||| 
Db         20 FNLDAEKLTHFHMDGAEFGHSVLQYDSSWVVVGAPKEIKATNQIGGLYKCGYHTGNCEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||:|||||||| |:|  |||||||| || || |: ||||  |: :|  |   
Db         80 SLQVPPEAVNISLGLSLAAATNPSWLLACGPTVHHTCRENIYLTGLCFLLSSSFKQS-QN 138

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              || | : ||::| || |||||||||    || :| ::|  || ||::  | |||||:|: | 
Db        139 FPTAQQECPKQDQDIVFLIDGSGSISSTDFEKMLDFVKAVMSQLQRPSTRFSLMQFSDYF 198

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |:|||  | :  :| ||:  : || | |:||: :: |: ||   :|||::| |:| ::  
Db        199 RVHFTFNNFISTSSPLSLLGSVRQLRGYTYTASAIKHVITELFTTQSGARQDATKVLIVI 258

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||| | || | |: |||  :   :||| |||| || :| |:|||  :|| |  ::|| :  
Db        259 TDGRKQGDNLSYDSVIPMAEAASIIRYAIGVGKAFYNEHSKQELKAIASMPSHEYVFSVE 318

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSPFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||:||| |:|||:|||||||||:| |||:|| |||||||||  | :||:|  |||: |:| 
Db        319 NFDALKDIENQLKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTPDGPVLGAVGSFSWSG 378

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | ||| |  : |||||:: : || ||||||: |:      | ||:||||||:|| | | :| 
Db        379 GAFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVVIFT 438

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   : |:||||||:|||||||||:| :|||||||| ||||| ||||||||||: 
Db        439 QESRHWRPKSEVRGTQIGSYFGASLCSVDMDRDGSTDLVLIGVPHYYEHTRGGQVSVCPM 498

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              | | :|| |   |:|||| |||||||||||||||| | ||||||||||:|||||||:| 
Db        499 P-GVGSRWHCGTTLHGEQGHPWGRFGAALTVLGDVNGDSLADVAIGAPGEEENRGAVYIF 557

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || |   |:|| |||| |::  |:|||||||||||||| |||||| ||::| |||||:: 
Db        558 HGASRQDIAPSPSQRISASQIPSRIQYFGQSLSGGQDLTRDGLVDLAVGSKGRVLLLRTR 617

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |:|||   : | | |::|:|||| :||  :   :  ||||: :| : :|  | ::| ||  
Db        618 PILRVSPTVHFTPAEISRSVFECQEQVAPEQTLSDATVCLHIHESPKTQL--GDLRSTVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              :|||| ||  :||:| |||      : : ||| : ||::|| || |:|| |:|| |||||  
Db        676 FDLALDHGRLSTRAIFKETKTRALTRVKTLGLNKHCESVKLLLPACVEDSVTPITLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||| |:|:  ||:|:|| | |  |||  ||||||| |:|||||||: | |  | |||| 
Db        736 SLVGVPISSLQNLQPMLAVDDQTYFTASLPFEKNCGADHICQDDLSVVFGFPDLKTLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQ----------NQRSQRSWR 829
              | || ||| ||||||| | || |:|: ||:|:|:  |            :| | |  
Db        796 SDLELNVDVTVSNDGEDSYGTTVTLFYPVGLSFRRVAEGQVFLRKKEDQQWQRRGQHSLH 855

Qy        830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSE 889
              | |:  |: : |  | ||||| | ||   |::|| :|||  || ||::|||:| | || 
Db        856 LMCD--STPDRSQGLWSTSCSSRHVIFRGGSQMTFLVTFDVSPKAELGDRLLLRARVGSE 913

Qy        890 NNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASE-NTSRVMQHQYQVSNLGQR 948
              ||:| | || ||||||||||||| :::||   ||||||: ||   : |::|::||:|||| 
Db        914 NNVPGTPKTTFQLELPVKYAVYTMISSHDQFTKYLNFSTSEKEKTSVVEHRFQVNNLGQR 973

Qy        949 SLPISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCS 1008
              :|:|: | ||: |   :|   |:  :| : |:      |:  | |  ::|:|||::|| 
Db        974 DVPVSINFWVPIELKGEAVW-TVMVSHPQNPLTQCYRNRLKPTQFDLLTHMQKSPVLDCS 1032
```

```
Qy       1009 IAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLP 1068
Db       1033 IADCLHLRCDIPSLGILDELYFILKGNLSFGWISQTLQKKVLLLSEAEITFNTSVYSQLP 1092

Qy       1069 GQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1093 GQEAFLRAQTKTVLEMYKVHNPVPLIVGSSVGGLLLLAIITAILYKAGFFKRQYKEMLEE 1152
```

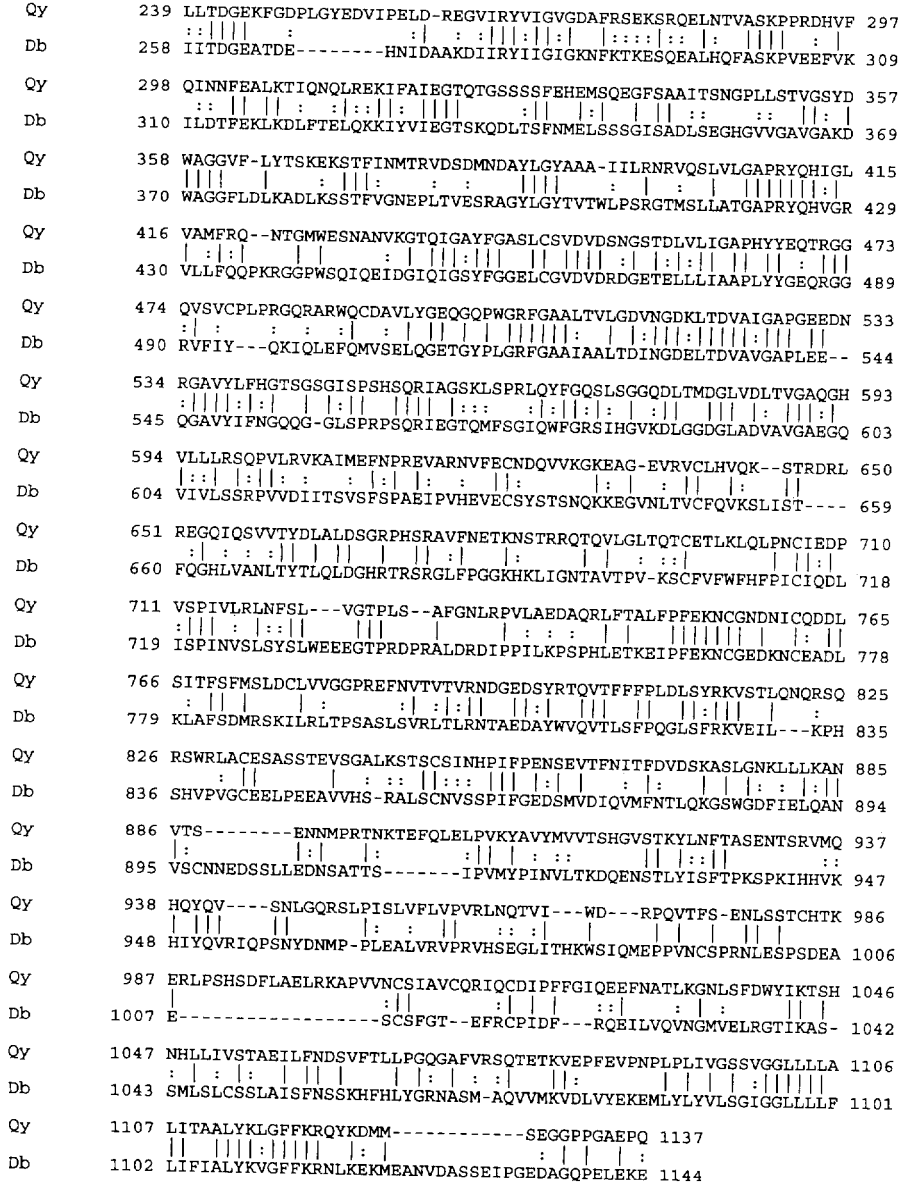
```
RESULT 10
ITAD_RAT
ID   ITAD_RAT       STANDARD;      PRT;  1161 AA.
AC   Q9QYE7;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-D precursor.
GN   Name=Itgad;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RA   O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA   Gallatin W.M.;
RT   "Cloning of rat alpha D, a novel beta 2 integrin.";
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC       VCAM1. May play a role in the atherosclerotic process such as
CC       clearing lipoproteins from plaques and in phagocytosis of blood-
CC       borne pathogens, particulate matter, and senescent erythrocytes
CC       from the blood (By similarity).
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC       associates with beta-2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF021334; AAF21241.1; -.
DR   HSSP; P11215; 1BHQ.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     19       Potential.
FT   CHAIN        20   1161       Integrin alpha-D.
FT   DOMAIN       20   1100       Extracellular (Potential).
FT   TRANSMEM   1101   1121       Potential.
FT   DOMAIN     1122   1161       Cytoplasmic (Potential).
FT   REPEAT       34     87       FG-GAP 1.
FT   REPEAT       88      ?       FG-GAP 2.
FT   DOMAIN      152    334       VWFA.
FT   REPEAT      352    402       FG-GAP 3.
FT   REPEAT      403    454       FG-GAP 4.
FT   REPEAT      456    517       FG-GAP 5.
FT   REPEAT      519    577       FG-GAP 6.
FT   REPEAT      582    634       FG-GAP 7.
FT   CA_BIND     467    475       Potential.
FT   CA_BIND     531    539       Potential.
FT   CA_BIND     594    602       Potential.
FT   SITE       1126   1130       GFFKR motif.
FT   DISULFID     69     76       By similarity.
FT   DISULFID    108    126       By similarity.
FT   DISULFID    656    711       By similarity.
FT   DISULFID    769    775       By similarity.
FT   DISULFID    845    860       By similarity.
FT   DISULFID    993   1017       By similarity.
FT   DISULFID   1022   1027       By similarity.
FT   CARBOHYD     61     61       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    245    245       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    393    393       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    696    696       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    734    734       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    784    784       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    907    907       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    936    936       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1045   1045       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   1161 AA;  126600 MW;  2258491A984A705E CRC64;
```

```
  Query Match             55.0%;  Score 3236.5;  DB 1;  Length 1161;
  Best Local Similarity   57.5%;  Pred. No. 1.6e-206;
  Matches  648;  Conservative  163;  Mismatches  303;  Indels   13;  Gaps    9;

Qy         2 NLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPIR 61
Db        21 NLDVEEPIVFREDAASFGQTVVQFGGSRLVVGAPLEAVAVNQTGRLYDCAPATGMCQPIV 80

Qy        62 LQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQKF 121
Db        81 LRSPLEAVNMSLGLSLVTATNNAQLLACGPTAQRACVKNMYAKGSCLLLGSSL-QFIQAV 139

Qy       122 PEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEFR 181
Db       140 PASMPECPRQEMDIAFLIDGSGSINQRDFAQMKDFVKALMGEFASTSTLFSLMQYSNILK 199

Qy       182 IHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLLT 241
Db       200 THFTFTEFKNILDPQSLVDPIVQLQGLTYTATGIRTVMEELFHSKNGSRKSAKKILLVIT 259

Qy       242 DGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQINN 301
Db       260 DGQKYRDPLEYSDVIPAADKAGIIRYAIGVGDAFQEPTALKELNTIGSAPPQDHVFKVGN 319

Qy       302 FEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAGG 361
Db       320 FAALRSIQRQLQEKIFAIEGTQSRSSSSFQHEMSQEGFSSALTSDGPVLGAVGSFSWSGG 379

Qy       362 VFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFRQ 421
Db       380 AFLYPPNTRPTFINMSQENVDMRDSYLGYSTAVAFWKGVHSLILGAPRHQHTGKVVIFTQ 439

Qy       422 NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPLP 481
Db       440 EARHWRPKSEVRGTQIGSYFGASLCSVDVDRDGSTDLVLIGAPHYYEQTRGGQVSVFPVP 499

Qy       482 RGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFH 541
Db       500 -GVRGRWQCEATLHGEQGHPWGRFGVALTVLGDVNGDNLADVAIGAPGEEESRGAVYIFH 558

Qy       542 GTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQP 601
Db       559 GASRLEIMPSPSQRVTGSQLSLRLQYFGQSLSGGQDLTQDGLVDLAVGAQGHVLLLRSLP 618
```

```
Qy        602 VLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVTY 661
              :|:|:  : | | |||: |::| ::      |||| ||| | | : | |  | :|  | |
Db        619 LLKVELSIRFAPMEVAKAVYQCWERTPTVLEAGEATVCLTVHKGSPDLL--GNVQGSVRY 676

Qy        662 DLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNFS 721
              ||||| ||  |||:|:|||| |    : |||   |||:|| ||:|:|| ||||:||||||
Db        677 DLALDPGRLISRAIFDETKNCTLTGRKTLGLGDHCETVKLLLPDCVEDAVSPIILRLNFS 736

Qy        722 LVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVGG 781
              ||    |   || |||| :|   ||  |||||| : :|: || |:|:|  |  |||||
Db        737 LVRDSASP-RNLHPVLAVGSQDHITASLPFEKNCKQELLCEGDLGISFNFSGLQVLVVGG 795

Qy        782 PREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEVS 841
              |  ||||| |:|||| | | |::|  ||||:|:  | |  |   |||||:  : :
Db        796 SPELTVTVTVWNEGEDSYGTLVKFYYPAGLSYRRVTGTQ-QPHQYPLRLACEAEPAAQED 854

Qy        842 GALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEFQ 901
              |:|:|||||||||| | :: || |||||  || ||::|||:|  :|||| | |||| ||
Db        855 --LRSSSCSINHPIFREGAKTTFMITFDVSYKAFLGDRLLLRAKASSENNKPDTNKTAFQ 912

Qy        902 LELPVKYAVYMVVTSHGVSTKYLNFTASENTSR-VMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||| || :::    || ::||::|    |    |:|:|:||    | : : | |||
Db        913 LELPVKYTVYTLISRQEDSTNHVNFSSSHGGRRQEAAHRYRVNNLSPLKLAVRVNFWVPV 972

Qy        961 RLNQTVIWDRPQVTFSENLSS-TCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              ||   :||  || |      :| :: : | : ||| ::::  |::|||| |   :|||
Db        973 LLNGVAVWD---VTLSSPAQGVSCVSQMKPPQNPDFLTQIQRRSVLDCSIADCLHFRCDI 1029

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              |   ||:| :  |:|||| |  :|    :|:|| ||| |: ||:: |||| ||:|:| |
Db       1030 PSLDIQDELDFILRGNLSFGWVSQTLQEKVLLVSEAEITFDTSVYSQLPGQEAFLRAQVE 1089

Qy       1080 TKVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMM 1126
              | :| : |  |: |: |||||||||||||  ||||||||||:|:
Db       1090 TTLEEYVVYEPIFLVAGSSVGGLLLLALITVVLYKLGFFKRQYKEML 1136
```

```
RESULT 11
ITAL_BOVIN
ID   ITAL_BOVIN     STANDARD;      PRT;  1165 AA.
AC   P61625;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE   alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE   chain) (CD11a).
GN   Name=ITGAL;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=14697514; DOI=10.1016/j.gene.2003.09.043;
RA   Fett T., Zecchinon L., Baise E., Desmecht D.;
RT   "The bovine (Bos taurus) CD11a-encoding cDNA: molecular cloning,
RT   characterisation and comparison with the human and murine
RT   glycoproteins.";
RL   Gene 325:97-101(2004).
CC   -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC       ICAM3 and ICAM4. It is involved in a variety of immune phenomena
CC       including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC       mediated killing, and antibody dependent killing by granulocytes
CC       and monocytes (By similarity).
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC       associates with beta-2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY267467; AAP94035.1; -.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL       1     23       Potential.
FT   CHAIN       24   1165       Integrin alpha-L.
FT   DOMAIN      24   1084       Extracellular (Potential).
FT   TRANSMEM  1085   1105       Potential.
FT   DOMAIN    1106   1165       Cytoplasmic (Potential).
FT   REPEAT      40     88       FG-GAP 1.
FT   REPEAT      89    146       FG-GAP 2.
FT   DOMAIN     153    324       VWFA.
FT   REPEAT     347    397       FG-GAP 3.
FT   REPEAT     398    452       FG-GAP 4.
FT   REPEAT     454    513       FG-GAP 5.
FT   REPEAT     515    572       FG-GAP 6.
FT   REPEAT     575    627       FG-GAP 7.
FT   CA_BIND    465    473       Potential.
FT   CA_BIND    527    535       Potential.
FT   CA_BIND    587    595       Potential.
FT   SITE      1111   1115       GFFKR motif.
FT   DISULFID    71     78       By similarity.
FT   DISULFID   108    126       By similarity.
FT   DISULFID   650    704       By similarity.
FT   DISULFID   768    774       By similarity.
FT   DISULFID   842    858       By similarity.
FT   DISULFID   994   1009       By similarity.
FT   DISULFID  1017   1048       By similarity.
FT   CARBOHYD    33     33       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    86     86       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   185    185       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   646    646       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   667    667       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   723    723       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   859    859       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   894    894       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   929    929       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1056   1056       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1067   1067       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE  1165 AA;  128725 MW;  DAEB3A3F1E1463CB CRC64;
```

```
  Query Match             26.7%;  Score 1572;  DB 1;  Length 1165;
  Best Local Similarity   34.6%;  Pred. No. 1.7e-95;
  Matches  409;  Conservative  201;  Mismatches  465;  Indels  108;  Gaps   33;

Qy          1 FNLDTENAMTFQENARG--FGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58
              :|||  :   |     |  ||  |:|:  |: |||||| |     |  |:||||   || |
Db         24 YNLDVRHVQNFSFPLAGRHFGYRVLQV-GNGVVVGAPSE---GNSMGNLYQCQPETGDCL 79

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQP 118
              |: |       :  ||::||   :   |||| | : :|| :| |: |||:|   |||
Db         80 PVTLS--SNYTSKYLGMTLATDPTSDNLLACDPGLSRTCDQNIYLSGLCYLIHENLRGPV 137

Qy        119 QKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSE 178
              :    : | : |: || ||| |:   :| :: :::  ||::|  |   |: : :|:|
Db        138 LQGHPGYQECIKGNVDLVFLFDGSMSLQQDEFEKIVDFMKDVMKKLSNSSYQFAAVQFST 197

Qy        179 EFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILF 238
              ||  ||| ::    :| :|:  :  :   |:|   :  | :|:|    ||| :| |:|
Db        198 YFRTEFTFLDYIRQKDPDALLAGVKHMRLLTNTFGAINYVAKEVFRPDLGARPDATKVLI 257
```



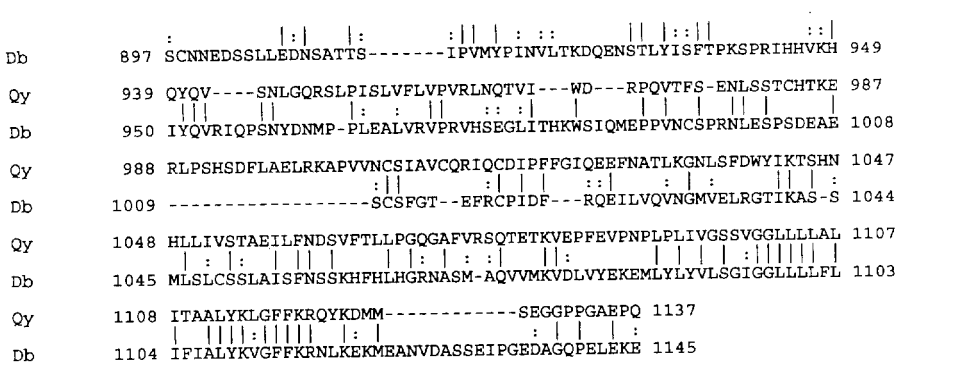
```
RESULT 12
ITAL_HUMAN
ID   ITAL_HUMAN     STANDARD;      PRT;  1170 AA.
AC   P20701; O43746;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE   alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE   chain) (CD11a).
GN   Name=ITGAL; Synonyms=CD11A;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RX   MEDLINE=89139587; PubMed=2537322;
RA   Larson R.S., Corbi A.L., Berman L., Springer T.;
RT   "Primary structure of the leukocyte function-associated molecule-1
RT   alpha subunit: an integrin with an embedded domain defining a protein
RT   superfamily.";
RL   J. Cell Biol. 108:703-712(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RX   MEDLINE=99425270; PubMed=10493829;
RA   Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
RA   Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA   Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
RA   Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT   "Genome duplications and other features in 12 Mb of DNA sequence from
RT   human chromosome 16p and 16q.";
RL   Genomics 60:295-308(1999).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.
RX   MEDLINE=96036067; PubMed=7479767;
RA   Qu A., Leahy D.J.;
RT   "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L
RT   beta 2) integrin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.
RX   MEDLINE=96398682; PubMed=8805579;
RA   Qu A., Leahy D.J.;
RT   "The role of the divalent cation in the structure of the I domain from
RT   the CD11a/CD18 integrin.";
RL   Structure 4:931-942(1996).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.
RX   MEDLINE=99425288; PubMed=10493852;
RA   Kallen J., Welzenbach K., Ramage P., Geyl D., Kriwacki R., Legge G.,
RA   Cottens S., Weitz-Schmidt G., Hommel U.;
RT   "Structural basis for LFA-1 inhibition upon lovastatin binding to the
RT   CD11a I-domain.";
RL   J. Mol. Biol. 292:1-9(1999).
CC   -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC       ICAM3 and ICAM4. It is involved in a variety of immune phenomena
CC       including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC       mediated killing, and antibody dependent killing by granulocytes
CC       and monocytes.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P20701-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P20701-2; Sequence=VSP_002738;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: Leukocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11a.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Y00796; CAA68747.1; -.
DR   EMBL; AC002310; AAC31672.1; -.
DR   PIR; S03308; S03308.
DR   PDB; 1CQP; X-ray; A/B=153-334.
DR   PDB; 1DGQ; NMR; A=149-336.
DR   PDB; 1LFA; X-ray; A/B=-.
DR   PDB; 1MJN; X-ray; A=153-331.
DR   PDB; 1MQ8; X-ray; B/D=155-331.
DR   PDB; 1MQ9; X-ray; A=152-330.
DR   PDB; 1MQA; X-ray; A=152-330.
DR   PDB; 1ZON; X-ray; @=150-336.
DR   PDB; 1ZOO; X-ray; A/B=150-336.
DR   PDB; 1ZOP; X-ray; A/B=150-336.
DR   Genew; HGNC:6148; ITGAL.
DR   MIM; 153370; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   3D-structure; Alternative splicing; Calcium; Cell adhesion;
KW   Direct protein sequencing; Glycoprotein; Integrin; Magnesium;
KW   Receptor; Repeat; Signal; Transmembrane.
FT   SIGNAL       1     25
FT   CHAIN       26   1170       Integrin alpha-L.
FT   DOMAIN      26   1090       Extracellular (Potential).
FT   TRANSMEM  1091   1111       Potential.
FT   DOMAIN    1112   1170       Cytoplasmic (Potential).
FT   REPEAT      42     91       FG-GAP 1.
FT   REPEAT      92    149       FG-GAP 2.
FT   DOMAIN     156    327       VWFA.
FT   REPEAT     350    400       FG-GAP 3.
FT   REPEAT     401    455       FG-GAP 4.
FT   REPEAT     457    516       FG-GAP 5.
FT   REPEAT     518    575       FG-GAP 6.
FT   REPEAT     578    630       FG-GAP 7.
FT   CA_BIND    468    476       Potential.
FT   CA_BIND    530    538       Potential.
FT   CA_BIND    590    598       Potential.
FT   SITE      1115   1119       GFFKR motif.
FT   DISULFID    73     80       By similarity.
FT   DISULFID   111    129       By similarity.
FT   DISULFID   653    707       By similarity.
FT   DISULFID   771    777       By similarity.
FT   DISULFID   845    861       By similarity.
FT   DISULFID   998   1013       By similarity.
FT   DISULFID  1021   1052       By similarity.
FT   CARBOHYD    65     65       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   188    188       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   649    649       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   670    670       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   726    726       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   730    730       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   862    862       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   885    885       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   897    897       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1060   1060       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1071   1071       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC   954    954       Q -> QGVHGLVEMQTSKQILCRPAGDAEHTVGAQEGELPC
FT                               PWGVSEAFRDNIRAGPCR (in isoform 2).
FT                               /FTId=VSP_002738.
FT   CONFLICT   214    214       R -> W (in Ref. 1 and 2).
FT   CONFLICT   660    660       Y -> I (in Ref. 2).
FT   STRAND     155    162
FT   STRAND     164    164
FT   TURN       165    166
FT   HELIX      169    185
FT   TURN       186    188
FT   STRAND     191    198
FT   STRAND     202    206
FT   HELIX      208    214
FT   HELIX      217    221
FT   TURN       222    223
FT   STRAND     229    229
FT   HELIX      233    243
FT   TURN       244    244
FT   HELIX      247    249
FT   TURN       250    250
FT   TURN       253    254
FT   STRAND     256    263
FT   HELIX      274    276
FT   TURN       277    278
FT   STRAND     280    286
FT   HELIX      288    290
FT   HELIX      293    297
FT   TURN       298    299
FT   HELIX      300    302
FT   HELIX      307    310
FT   STRAND     311    314
FT   HELIX      317    319
FT   HELIX      323    330
SQ   SEQUENCE  1170 AA;  128819 MW;  39A7AF92EF286FC0 CRC64;
```

Query Match 26.3%; Score 1547.5; DB 1; Length 1170;
Best Local Similarity 34.3%; Pred. No. 7.2e-94;
Matches 403; Conservative 210; Mismatches 460; Indels 101; Gaps 37;

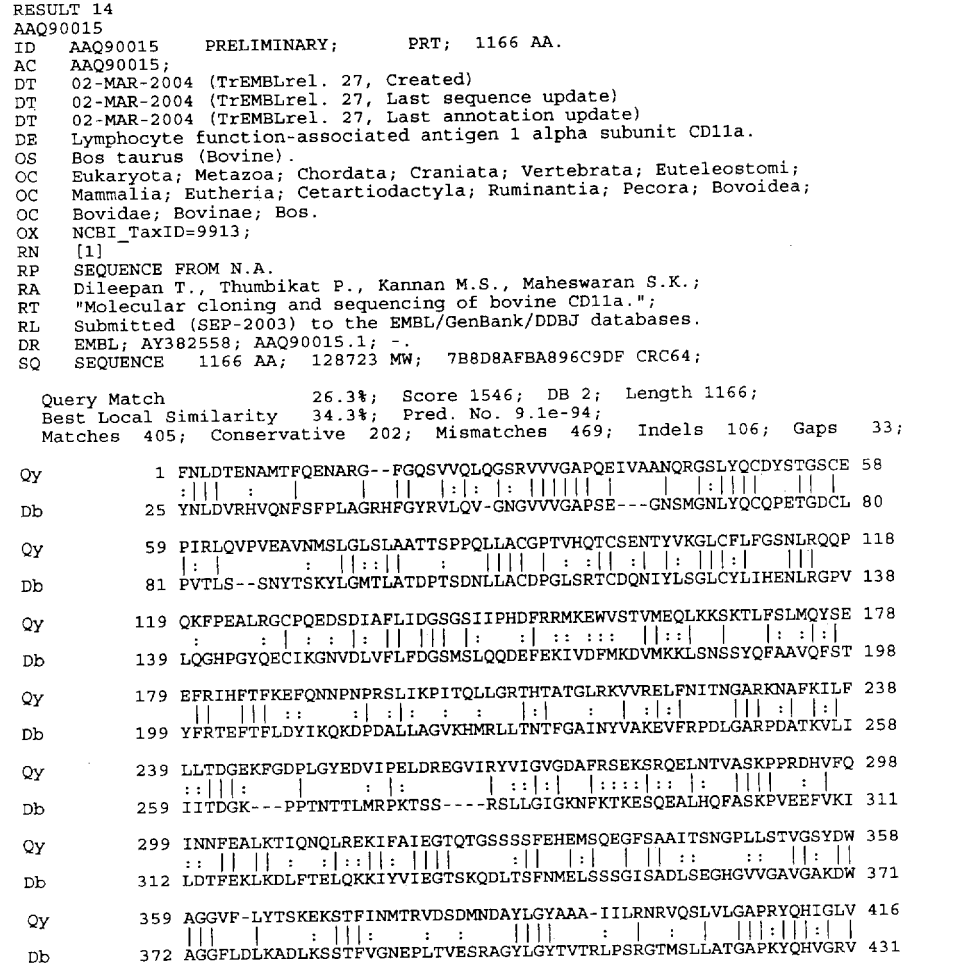
```
Qy          1 FNLDTENAMTFQ--ENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58
              :|||   | :|      | ||  |:|: |: |:|||| |    | ||||||   || |
Db         26 YNLDVRGARSFSPPRAGRHFGYRVLQV-GNGVIVGAPGE---GNSTGSLYQCQSGTGHCL 81

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLR--- 115
              |: |:      :  ||::||   :   :||| | : :|| :|||: |||:|| ||:
Db         82 PVTLR-GSNYTSKYLGMTLATDPTDGSILACDPGLSRTCDQNTYLSGLCYLFRQNLQGPM 140

Qy        116 -QQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLM 174
              |   | | ::|    : |: || ||| |: | :|::: :::  ||::|  :   |: :
Db        141 LQGRPGFQECIKG----NVDLVFLFDGSMSLQPDEFQKILDFMKDVMKKLSNTSYQFAAV 196

Qy        175 QYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAF 234
              |:|  ::  | | ::    :| :|:| :  :|  |:|   :  |  |:|    ||| :|
Db        197 QFSTSYKTEFDFSDYVKRKDPDALLKHVKHMLLLTNTFGAINYVATEVFREELGARPDAT 256

Qy        235 KILFLLTDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRD 294
              |:| ::|||||      |  |   :     :|||:||:|  |::::|:: |:  |||| :
Db        257 KVLIIITDGE--ATDSGNIDAAKD-----IIRYIIGIGKHFQTKESQETLHKFASKPASE 309

Qy        295 HVFQINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVG 354
              |  :: || || :  :|::||: ||||      :||  |:|  | || ::    ::  ||
Db        310 FVKILDTFEKLKDLFTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSRGHAVVGAVG 369

Qy        355 SYDWAGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQH 412
              : ||||| |   :  |||   :  ::   ||||    :  | :   |  |||||||
Db        370 AKDWAGGFLDLKADLQDDTFIGNEPLTPEVRAGYLGYTVTWLPSRQKTSLLASGAPRYQH 429

Qy        413 IGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQT 470
              :| | :|:  |  | |     : |||||:|||  || |||| :|  |:|:|||| :| :
Db        430 MGRVLLFQEPQGGGHWSQVQTIHGTQIGSYFGGELCGVDVDQDGETELLLIGAPLFYGEQ 489

Qy        471 RGGQVSVCPLPRGQRARWQCDAV--LYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAP 528
              |||:| :      || :   : |  | |: | | |||| |:| | |:||| | |||:|||
```

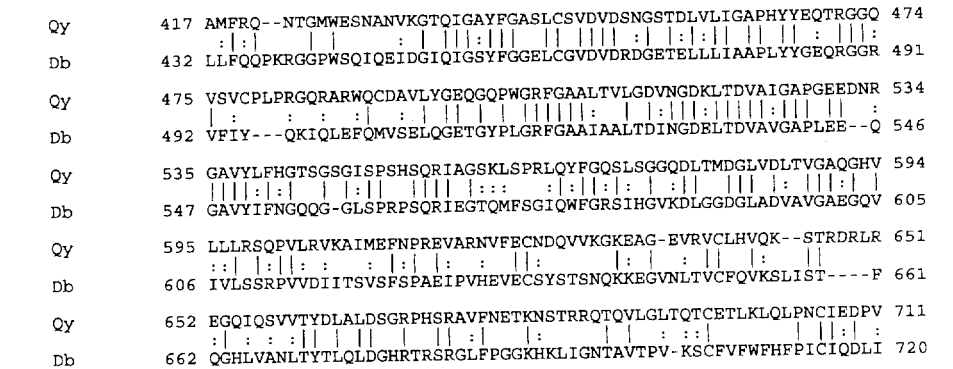
```
Db        490 RGGRVFIY-----QRRQLGFEEVSELQGDPGYPLGRFGEAITALTDINGDGLVDVAVGAP 544

Qy        529 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTV 588

Db        545 LEE--QGAVYIFNGRHG-GLSPQPSQRIEGTQVLSGIQWFGRSIHGVKDLEGDGLADVAV 601

Qy        589 GAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVV-KGKEAGEVRVCLHVQKSTR 647

Db        602 GAESQMIVLSSRPVVDMVTLMSFSPAEIPVHEVECSYSTSNKMKEGVNITICFQI-KSLY 660

Qy        648 DRLREGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCI 707

Db        661 PQF-QGRLVANLTYTLQLDGHRTRRRGLFPGGRHELRRNIAVT-TSMSCTDFSFHFPVCV 718

Qy        708 EDPVSPIVLRLNFSL---VGTPLS--AFGN-----LRPVLAEDAQRLFTALFPFEKNCGN 757

Db        719 QDLISPINVSLNFSLWEEEGTPRDQRAQGKDIPPILRPSLHSETWEI-----PFEKNCGE 773

Qy        758 DNICQDDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVS 817

Db        774 DKKCEANLRVSFSPARSRALRLTAFASLSVELSLSNLEEDAYWVQLDLHFPPGLSFRKVE 833

Qy        818 TLQNQRSQRSWRLACES--ASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKAS 875

Db        834 ML---KPHSQIPVSCEELPEESRLLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNSS 887

Qy        876 LGNKLLLKANVTSENN----MPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASEN 931

Db        888 WGDSVELHANVTCNNEDSDLLEDNSATTI---IPILYPINILIQDQEDSTLYVSFTPKGP 944

Qy        932 TSRVMQHQYQV---SNLGQRSLP-ISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTK- 986

Db        945 KIHQVKHMYQVRIQPSIHDHNIPTLEAVVGVPQPPSEGPITHQWSVQMEPPV--PCHYED 1002

Qy        987 -ERLPSHSD--FLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIK 1043

Db       1003 LERLPDAAEPCLPGALFRCPVV----------------FRQEILVQVIGTLELVGEIE 1044

Qy       1044 TSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLL 1103

Db       1045 AS-SMFSLCSSLSISFNSSKHFHLYGSNASL-AQVVMKVDVVYEKQMLYLYVLSGIGGLL 1102

Qy       1104 LLALITAALYKLGFFKRQYKDMMSEG-GPPGAEP 1136

Db       1103 LLLLIFIVLYKVGFFKRNLKEKMEAGRGVPNGIP 1136


RESULT 13
Q6TYB8
ID   Q6TYB8      PRELIMINARY;      PRT;  1166 AA.
AC   Q6TYB8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR   EMBL; AY382558; AAQ90015.1; -.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 1.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Cell adhesion; Integrin; Transmembrane.
SQ   SEQUENCE   1166 AA;  128723 MW;  7B8D8AFBA896C9DF CRC64;

  Query Match             26.3%;  Score 1546;  DB 2;  Length 1166;
  Best Local Similarity   34.3%;  Pred. No. 9.1e-94;
  Matches  405;  Conservative  202;  Mismatches  469;  Indels  106;  Gaps   33;

Qy          1 FNLDTENAMTFQENARG--FGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58

Db         25 YNLDVRHVQNFSFPLAGRHFGYRVLQV-GNGVVVGAPSE---GNSMGNLYQCQPETGDCL 80

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQP 118

Db         81 PVTLS--SNYTSKYLGMTLATDPTSDNLLACDPGLSRTCDQNIYLSGLCYLIHENLRGPV 138

Qy        119 QKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSE 178

Db        139 LQGHPGYQECIKGNVDLVFLFDGSMSLQQDEFEKIVDFMKDVMKKLSNSSYQFAAVQFST 198

Qy        179 EFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILF 238

Db        199 YFRTEFTFLDYIKQKDPDALLAGVKHMRLLTNTFGAINYVAKEVFRPDLGARPDATKVLI 258

Qy        239 LLTDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQ 298

Db        259 IITDGK---PPTNTTLMRPKTSS----RSLLGIGKNFKTKESQEALHQFASKPVEEFVKI 311

Qy        299 INNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDW 358

Db        312 LDTFEKLKDLFTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSEGHGVVGAVGAKDW 371

Qy        359 AGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQHIGLV 416

Db        372 AGGFLDLKADLKSSTFVGNEPLTVESRAGYLGYTVTRLPSRGTMSLLATGAPKYQHVGRV 431

Qy        417 AMFRQ--NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQ 474

Db        432 LLFQQPKRGGPWSQIQEIDGIQIGSYFGGELCGVDVDRDGETELLLIAAPLYYGEQRGGR 491

Qy        475 VSVCPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNR 534

Db        492 VFIY---QKIQLEFQMVSELQGETGYPLGRFGAAIAALTDINGDELTDVAVGAPLEE--Q 546

Qy        535 GAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHV 594

Db        547 GAVYIFNGQQG-GLSPRPSQRIEGTQMFSGIQWFGRSIHGVKDLGGDGLADVAVGAEGQV 605

Qy        595 LLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAG-EVRVCLHVQK--STRDRLR 651

Db        606 IVLSSRPVVDIITSVSFSPAEIPVHEVECSYSTSNQKKEGVNLTVCFQVKSLIST----F 661

Qy        652 EGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPV 711

Db        662 QGHLVANLTYTLQLDGHRTRSRGLFPGGKHKLIGNTAVTPV-KSCFVFWFHFPICIQDLI 720

Qy        712 SPIVLRLNFSL---VGTPLS--AFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLS 766

Db        721 SPINVSLSYSLWEEEGTPRDPRALDRDIPPILKPSPHLETKEIPFEKNCGEDKNCEADLK 780

Qy        767 ITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQR 826

Db        781 LAFSDMRSKILRLTPSASLSVRLTLRNTAEDAYWVQVTLSFPQGLSFRKVEIL---KPHS 837

Qy        827 SWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANV 886

Db        838 HVPVGCEELPEEAVVHS-RALSCNVSSPIFGEDSMVDIQVMFNTLQKGSWGDFIELQANV 896

Qy        887 TS--------ENNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQH 938
```



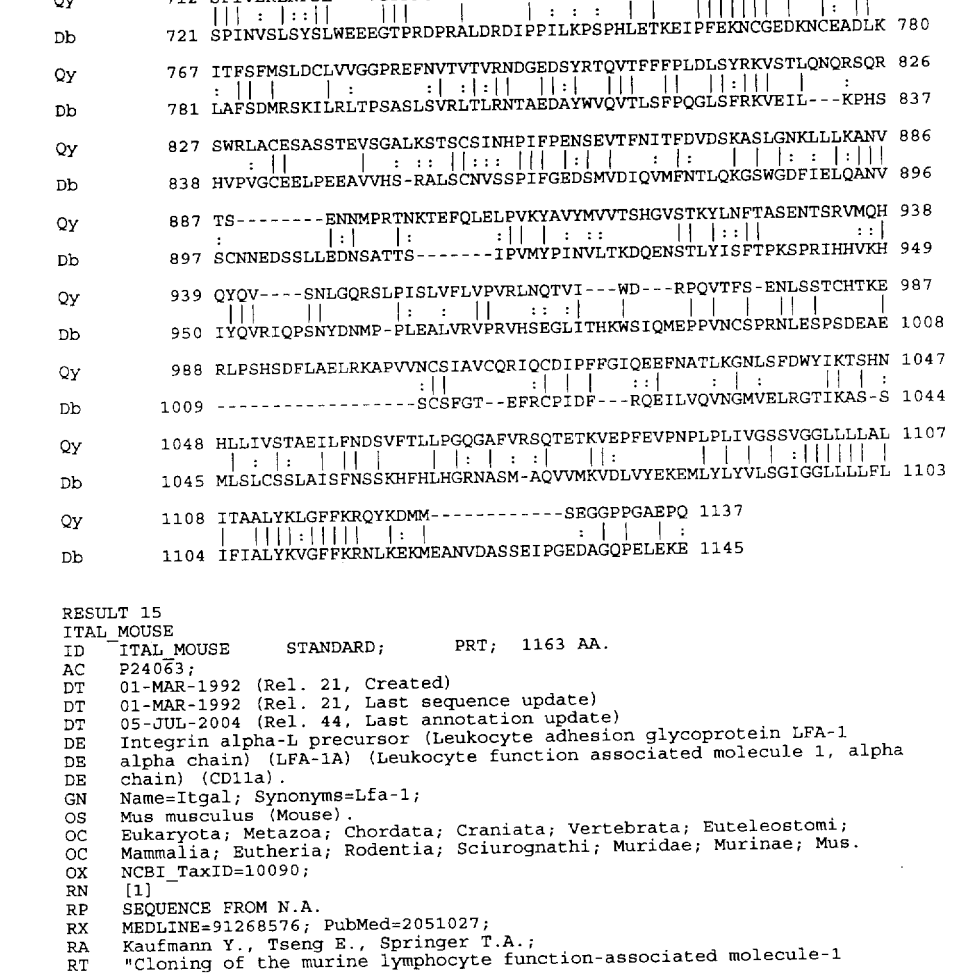
RESULT 14
AAQ90015
ID AAQ90015 PRELIMINARY; PRT; 1166 AA.
AC AAQ90015;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT "Molecular cloning and sequencing of bovine CD11a.";
RL Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY382558; AAQ90015.1; -.
SQ SEQUENCE 1166 AA; 128723 MW; 7B8D8AFBA896C9DF CRC64;

Query Match 26.3%; Score 1546; DB 2; Length 1166;
Best Local Similarity 34.3%; Pred. No. 9.1e-94;
Matches 405; Conservative 202; Mismatches 469; Indels 106; Gaps 33;





RESULT 15
ITAL_MOUSE
ID ITAL_MOUSE STANDARD; PRT; 1163 AA.
AC P24063;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE chain) (CD11a).
GN Name=Itgal; Synonyms=Lfa-1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91268576; PubMed=2051027;
RA Kaufmann Y., Tseng E., Springer T.A.;
RT "Cloning of the murine lymphocyte function-associated molecule-1



```
RT   alpha-subunit and its expression in COS cells.";
RL   J. Immunol. 147:369-374(1991).
RN   [2]
RP   SEQUENCE OF 24-42.
RX   MEDLINE=85188276; PubMed=3887182;
RA   Springer T.A., Teplow D.B., Dreyer W.J.;
RT   "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
RT   glycoproteins and unexpected relation to leukocyte interferon.";
RL   Nature 314:540-542(1985).
CC   -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC       ICAM3 and ICAM4. Is involved in a variety of immune phenomena
CC       including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC       mediated killing, and antibody dependent killing by granulocytes
CC       and monocytes. Mice expressing a null mutation of the alpha-L
CC       subunit gene demonstrate impaired tumor rejection and impaired
CC       leukocytes recruitment.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Leukocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M60778; AAA39426.1; -.
DR   PIR; I56126; I56126.
DR   HSSP; P20701; 1DGQ.
DR   MGD; MGI:96606; Itgal.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 2.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Direct protein sequencing; Glycoprotein;
KW   Integrin; Magnesium; Receptor; Repeat; Signal; Transmembrane.
FT   SIGNAL        1     23
FT   CHAIN        24   1163       Integrin alpha-L.
FT   DOMAIN       24   1087       Extracellular (Potential).
FT   TRANSMEM   1088   1108       Potential.
FT   DOMAIN     1109   1163       Cytoplasmic (Potential).
FT   REPEAT       39     88       FG-GAP 1.
FT   REPEAT       89    146       FG-GAP 2.
FT   DOMAIN      153    325       VWFA.
FT   REPEAT      348    398       FG-GAP 3.
FT   REPEAT      399    454       FG-GAP 4.
FT   REPEAT      455    514       FG-GAP 5.
FT   REPEAT      516    573       FG-GAP 6.
FT   REPEAT      576    628       FG-GAP 7.
FT   CA_BIND     466    474       Potential.
FT   CA_BIND     528    536       Potential.
FT   CA_BIND     588    596       Potential.
FT   SITE       1111   1115       GFFKR motif.
FT   DISULFID     70     77       By similarity.
FT   DISULFID    108    126       By similarity.
FT   DISULFID    147    199       By similarity.
FT   DISULFID    651    705       By similarity.
FT   DISULFID    767    773       By similarity.
FT   DISULFID    840    856       By similarity.
FT   DISULFID    993   1009       By similarity.
FT   DISULFID   1017   1048       By similarity.
FT   CARBOHYD     86     86       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    185    185       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    270    270       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    444    444       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    668    668       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    696    696       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    724    724       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    728    728       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    776    776       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    857    857       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    880    880       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    890    890       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    899    899       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    927    927       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1056   1056       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   1163 AA;  128343 MW;  A7A3078489E8232F CRC64;

  Query Match             26.0%;  Score 1527.5;  DB 1;  Length 1163;
  Best Local Similarity   34.1%;  Pred. No. 1.5e-92;
  Matches  398;  Conservative  217;  Mismatches  463;  Indels   89;  Gaps   36;

Qy          1 FNLDTENAMTFQENA-RGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEP 59
              :||||    :|   | | ||  |:|::   ||||||:|    :  | || |  |:  |:|
Db         24 YNLDTRPTQSFLAQAGRHFGYQVLQIEDG-VVVGAPGE---GDNTGGLYHCRTSSEFCQP 79

Qy         60 IRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQ 119
              : |       :  ||::||    :   |||| | : :|| :|||: |||:||  :|
Db         80 VSLH-GSNHTSKYLGMTLATDAAKGSLLACDPGLSRTCDQNTYLSGLCYLFPQSLEGPML 138

Qy        120 KFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEE 179
              :   | : | :   |: || ||| |:   || :: |::  || :|  :   |: :|:| :
Db        139 QNRPAYQECMKGKVDLVFLFDGSQSLDRKDFEKILEFMKDVMRKLSNTSYQFAAVQFSTD 198

Qy        180 FRIHFTFKEF-QNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILF 238
              |  ||| :: : | ||  |:  :  :   |:|   :  ||  :|   :||| :| |:|
Db        199 CRTEFTFLDYVKQNKNPDVLLGSVQPMFLLTNTFRAINYVVAHVFKEESGARPDATKVLV 258

Qy        239 LLTDGEKF--GDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHV 296
              ::||||    |:     |          : ||:||:|  | | : :: |:  ||:|  : |
Db        259 IITDGEASDKGNISAAHD---------ITRYIIGIGKHFVSVQKQKTLHIFASEPVEEFV 309

Qy        297 FQINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSY 356
              :: || || :   |: :|:|||||      :||  |:|  |  || ::    ::  ||:
Db        310 KILDTFEKLKDLFTDLQRRIYAIEGTNRQDLTSFNMELSSSGISADLSKGHAVVGAVGAK 369

Qy        357 DWAGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQHIG 414
              |||||   |    : :||:    : ||:   ||||  | :  |:    |  |||||||:|
Db        370 DWAGGFLDLREDLQGATFVGQEPLTSDVRGGYLGYTVAWMTSRSSRPLLAAGAPRYQHVG 429

Qy        415 LVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRG 472
              | :|:  :  | |      ::|||||:|||  |||||:| :|  :|:|||||  :: : ||
Db        430 QVLLFQAPEAGGRWNQTQKIEGTQIGSYFGGELCSVDLDQDGEAELLLIGAPLFFGEQRG 489

Qy        473 GQVSVCPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEED 532
              |:|      : ::: ::  : | |: | | ||||||:|  | |:|||:|||||:||| ||
Db        490 GRVFTY---QRRQSLFEMVSELQGDPGYPLGRFGAAITALTDINGDRLTDVAVGAPLEE- 545

Qy        533 NRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQG 592
              :||||:|:|  | |:||  |||| |::: | :::||:|: | :||  | / |: |||:|
Db        546 -QGAVYIFNGKPG-GLSPQPSQRIQGAQVFPGIRWFGRSIHGVKDLGGDRLADVVVGAEG 603

Qy        593 HVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAG-EVRVCLHVQKSTRDRLR 651
              |::| |:||: |   : |:| |:  :   ||:       : :: | ::: |  ::  |
Db        604 RVVVLSSRPVVDVVTELSFSPEEIPVHEVECSYSAREEQKHGVKLKACFRIKPLTPQ--F 661

Qy        652 EGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPV 711
              :|:: : ::| | ||  |  || :| :  :     | :    ::|       | ||:| :
Db        662 QGRLLANLSYTLQLDGHRMRSRGLFPDGSHELSGNTSITP-DKSCLDFHFHFPICIQDLI 720

Qy        712 SPIVLRLNFSLV---GTPLSAFGN-LRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSI 767
```

```
                |||  : |||||:   |||     |  ::|:|       : |   |||||||  |  |: :|::
Db        721 SPINVSLNFSLLEEEGTPRDQKGRAMQPILRPSIHTV-TKEIPFEKNCGEDKKCEANLTL 779

Qy        768 TFSFMSLDCLVVGGP------REFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQN 821
                :    |        ||           |  |: | |||:|  ::    ||  ||:|||  ||
Db        780 SSPARS-------GPLRLMSSASLAVEWTLSNSGEDAYWVRLDLDFPRGLSFRKVEMLQ- 831

Qy        822 QRSQRSWRLACESASSTEVSGAL-KSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKL 880
                    ::||    || |  | |:  |::: |||    ||:  : |:    :|  : :
Db        832 --PHSRMPVSCEEL--TEGSSLLTKTLKCNVSSPIFKAGQEVSLQVMFNTLLNSSWEDFV 887

Qy        881 LLKANVTSEN-NMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQ 939
                |   |  || |                :|| | | ::      || |::||     :: :||
Db        888 ELNGTVHCENENSSLQEDNSAATHIPVLYPVNILTKEQENSTLYISFTPKGPKTQQVQHV 947

Qy        940 YQVSNLGQRSLPISLVFLVPVRLNQTVIWDRPQ----VTFSENLSS----TCHTKE-RLP 990
                |||      |  | :    :|  |   |   ||     :|:: :: :     |||::: : |
Db        948 YQV-----RIQPSAYDHNMPT-LEALVGVPRPHSEDLITYTWSVQTDPLVTCHSEDLKRP 1001

Qy        991 SHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLL 1050
                |      ||     | |        : :| | |   : |    : | :     || | : |
Db       1002 SSE---AEQPCLPGV--------QFRCPIVF---RWEILIQVTGTVELSKEIKAS-STLS 1046

Qy       1051 IVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLALITA 1110
                : |:  : || |    | |  |   :|   ||:       | : | | :|||:|| ||
Db       1047 LCSSLSVSFNSSKHFHLYGSKA-SEAQVLVKVDLIHEKEMLHVYVLSGIGGLVLLFLIFL 1105

Qy       1111 ALYKLGFFKRQYKDMM-SEGGPPGAEP 1136
                ||||:|||||  |: | ::|| |   |
Db       1106 ALYKVGFFKRNLKEKMEADGGVPNGSP 1132
```

Search completed: November 9, 2004, 12:17:33
Job time : 142.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:45:42 ; Search time 103.25 Seconds
(without alignments)
3950.365 Million cell updates/sec

Title: US-09-902-481B-4
Perfect score: 5884
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5852	99.5	1153	2	AAW65090	Aaw65090 Human Bet
2	5852	99.5	1153	3	AAB07360	Aab07360 Human CD1
3	5852	99.5	1153	5	AAU80252	Aau80252 Human int
4	5852	99.5	1153	5	ABG61469	Abg61469 Human Bet
5	5852	99.5	1153	5	AAO14428	Aao14428 Integrin
6	5852	99.5	1153	7	ADD25615	Add25615 Binding d
7	5842	99.3	1153	2	AAR04136	Aar04136 Alpha sub
8	5836.5	99.2	1152	8	ADM99589	Adm99589 Human int
9	5829.5	99.1	1152	8	ADP12435	Adp12435 Protein e
10	3473	59.0	1163	8	ADP44061	Adp44061 Human CD1
11	3459	58.8	1163	8	ADN02004	Adn02004 Human inf
12	3459	58.8	1163	8	ADQ17510	Adq17510 Human sof
13	3450	58.6	1163	2	AAR07120	Aar07120 p150.95 a
14	3436	58.4	1163	2	AAW65091	Aaw65091 Human Bet
15	3436	58.4	1163	3	AAB07361	Aab07361 Human CD1
16	3436	58.4	1163	5	ABG61470	Abg61470 Human Bet
17	3434	58.4	1163	6	ABU07406	Abu07406 Protein d
18	3434	58.4	1163	7	ADG32005	Adg32005 Human hom
19	3401	57.8	1161	2	AAR78166	Aar78166 Human bet
20	3401	57.8	1161	2	AAW23049	Aaw23049 Human bet
21	3401	57.8	1161	2	AAW57491	Aaw57491 Human bet
22	3401	57.8	1161	2	AAW65089	Aaw65089 Human Bet
23	3401	57.8	1161	2	AAW72825	Aaw72825 Human alp
24	3401	57.8	1161	2	AAW73342	Aaw73342 Human alp
25	3401	57.8	1161	3	AAB07359	Aab07359 Human alp
26	3401	57.8	1161	5	ABG61468	Abg61468 Human Bet
27	3385.5	57.5	1161	2	AAW23064	Aaw23064 Human bet
28	3385.5	57.5	1161	2	AAW65106	Aaw65106 Human Bet
29	3385.5	57.5	1161	2	AAW72837	Aaw72837 Human alp
30	3385.5	57.5	1161	2	AAW73343	Aaw73343 Human alp
31	3385.5	57.5	1161	3	AAB07376	Aab07376 Human alp
32	3385.5	57.5	1161	5	ABG61485	Abg61485 Human Bet
33	3226.5	54.8	1161	2	AAR78169	Aar78169 Rat alpha
34	3224.5	54.8	1161	2	AAW23062	Aaw23062 Rat beta
35	3224.5	54.8	1161	2	AAW60004	Aaw60004 Rat alpha
36	3224.5	54.8	1161	2	AAW72824	Aaw72824 Rat alpha
37	3224.5	54.8	1161	3	AAB07374	Aab07374 Rat alpha
38	3224.5	54.8	1161	5	ABG61483	Abg61483 Rat Beta2
39	3217.5	54.7	1161	2	AAW65104	Aaw65104 Rat beta-
40	3217.5	54.7	1161	2	AAW73345	Aaw73345 Rat alpha
41	3212	54.6	1161	2	AAW23061	Aaw23061 Mouse bet
42	3212	54.6	1161	2	AAW60003	Aaw60003 Mouse alp
43	3212	54.6	1161	2	AAW65103	Aaw65103 Mouse bet
44	3212	54.6	1161	2	AAW72836	Aaw72836 Mouse alp
45	3212	54.6	1161	2	AAW73347	Aaw73347 Mouse alp

ALIGNMENTS

```
RESULT 1
AAW65090
ID   AAW65090 standard; protein; 1153 AA.
XX
AC   AAW65090;
XX
DT   28-SEP-1998  (first entry)
XX
DE   Human Beta-integrin CD11b subunit protein.
XX
KW   Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW   type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
KW   lung inflammation; acute respiratory distress syndrome; CD11b subunit;
KW   rheumatoid arthritis.
XX
OS   Homo sapiens.
XX
PN   US5728533-A.
XX
PD   17-MAR-1998.
XX
PF   07-JUN-1995;   95US-00485618.
XX
PR   23-DEC-1993;   93US-00173497.
PR   05-AUG-1994;   94US-00286889.
PR   21-DEC-1994;   94US-00362652.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Van Der Vieren M,  Gallatin WM;
XX
DR   WPI; 1998-206565/18.
XX
PT   Screening assay for modulators of integrin binding - using immobilised or
PT   labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.
XX
PS   Example 5; Fig 1A-D; 106pp; English.
XX
CC   This sequence represents a human beta-integrin CD11b subunit which is
CC   used to describe a method for identifying compounds that modulate the
CC   interaction of the beta-integrin alpha-d subunit with a binding partner
CC   of alpha-d which involves contacting an alpha-d polypeptide with an alpha
CC   -d binding partner, one of which is immobilised and the other of which is
CC   labelled, in the presence of a test compound, and determining if the
CC   compound affects binding between the alpha-d polypeptide and alpha-d
CC   binding partner, where the alpha-d polypeptide is alpha-d or its fragment
CC   comprising the cytoplasmic, transmembrane or extracellular domain of
CC   alpha-d. Compounds that modulate alpha-d binding could be used to treat
```

```
CC   diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC   asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC   and rheumatoid arthritis
XX
SQ   Sequence 1153 AA;

  Query Match             99.5%;  Score 5852;  DB 2;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:||||||||||:||||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| ||||||||||||||||||||||||:|||||||||||:|||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 2
AAB07360
ID   AAB07360 standard; protein; 1153 AA.
XX
AC   AAB07360;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Human CD11b protein sequence.
XX
KW   Human; macrophage infiltration inhibition; alpha_d integrin;
KW   leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW   inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW   atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW   lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW   rheumatoid arthritis; central nervous system injury; CD11b.
XX
OS   Homo sapiens.
XX
PN   WO200029446-A1.
XX
PD   25-MAY-2000.
XX
PF   16-NOV-1999;   99WO-US027139.
XX
PR   16-NOV-1998;   98US-00193043.
PR   08-JUL-1999;   99US-00350259.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin MW,  Van Der Vieren M;
XX
DR   WPI; 2000-387751/33.
XX
PT   Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT   macrophage infiltration and reduce inflammation at central nervous system
PT   injury sites.
XX
PS   Example 5; Fig 1; 270pp; English.
XX
CC   Integrins are a class of membrane-associated molecules that participate
CC   in cellular adhesion. Integrins are made up of an alpha subunit and a
CC   beta subunit. One class of human integrins are restricted to expression
CC   in white blood cells and have a common beta2 subunit: the leukocyte
CC   integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC   have an important role in immune and inflammatory responses. The present
CC   protein sequence is the human integrin alpha subunit CD11b. This sequence
CC   was used in an alignment to identify a novel beta2 integrin alpha
CC   subunit: alpha_d (AAA60014 and AAB07359). The present sequence has
CC   approximately 60% identity to the protein sequence of alpha_d. The
CC   Alpha_d gene and protein may be useful in therapy for diseases linked to
CC   aberrant alpha_d function e.g. Type I diabetes, atherosclerosis, multiple
CC   sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
CC   distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
CC   (LAD). In addition, anti-alpha_d monoclonal antibodies may be used in the
CC   inhibition of macrophage infiltration at the site of a central nervous
CC   system injury. The monoclonal antibodies can also be used to detect and
CC   diagnose Crohn's disease
XX
```

SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 3; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:|||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||||:||||||||||:|:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 3
AAU80252
ID AAU80252 standard; protein; 1153 AA.
XX
AC AAU80252;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human integrin 1 alpha-M subunit protein.
XX
KW Integrin; antiinflammatory; imunosuppression; nephritis; dermatitis;
KW inflammatory disease; autoimmune disorder; Crohn's disease;
KW human immunodeficiency virus; HIV; myocardial infarction;
KW Sjorgen's syndrome; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 499. .500
FT /note= "Encoded by GGG CAG AGG"
XX
PN WO200218583-A2.
XX
PD 07-MAR-2002.
XX
PF 31-AUG-2001; 2001WO-US027227.
XX
PR 01-SEP-2000; 2000US-0229700P.
XX
PA (BLOO-) CENT BLOOD RES INC.
XX
PI Springer TA, Shimoaka M, Lu C;
XX
DR WPI; 2002-382964/41.
DR N-PSDB; ABK50046.
XX
PT Modified integrin-I or integrin I-like domain polypeptide useful as an
PT immunogen to produce antibodies specific to polypeptide, comprises a
PT disulfide bond such that polypeptide is stabilized in a desired
PT conformation.
XX
PS Disclosure; Page 109-112; 112pp; English.
XX
CC This invention relates to a modified integrin-I or integrin I-like domain
CC polypeptide comprising at least one disulfide bond so that the domain is
CC stabilised in a desired conformation. The polypeptide of the invention
CC may have antiinflammatory or imunosuppressive activities. The
CC polypeptides of the invention have an open conformation and are useful as
CC immunogens to produce antibodies that selectively bind to integrin I-
CC domain; and for identifying a modulator of integrin activity, or of
CC interaction of an integrin and a cognate ligand. The polypeptide of the
CC invention, or antibodies (preferably anti-LFA-1 antibody) is useful for
CC treating or preventing an integrin mediated disorder which is an
CC inflammatory or autoimmune disorder in a subject and for inhibiting the
CC binding of an integrin to a cognate ligand such as Crohn's disease,
CC nephritis; human immunodeficiency virus (HIV), myocardial infarction,
CC Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic
CC composition comprising the peptide of the invention is useful for
CC treating an integrin mediated disorder in a subject. The polypeptides
CC and/or active or antigenic fragments are useful as reagents for diagnosis
CC of integrin-mediated disorders. The present sequence represents the human
CC integrin-1 alpha-M protein subunit used to generate the mutant
CC polypeptides of the invention

```
XX
SQ   Sequence 1153 AA;

  Query Match            99.5%;  Score 5852;  DB 5;  Length 1153;
  Best Local Similarity  99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| ||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 4
ABG61469
ID   ABG61469 standard; protein; 1153 AA.
XX
AC   ABG61469;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human Beta2 integrin alphaCD11b subunit.
XX
KW   Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; LAD;
KW   leukocyte adhesion deficiency; inflammatory response; diabetes;
KW   multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW   inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW   immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
KW   intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW   locomotor recovery; locomotor damage; locomotor impairment;
KW   autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX
OS   Homo sapiens.
XX
PN   WO200230980-A2.
XX
PD   18-APR-2002.
XX
PF   15-OCT-2001; 2001WO-US032059.
XX
PR   13-OCT-2000; 2000US-00688307.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin WM,  Van Der Vieren M;
XX
DR   WPI; 2002-463260/49.
XX
PT   Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT   recovery, inhibiting locomotor damage, limiting locomotor impairment, or
PT   limiting autonomic and sensory dysfunction following spinal cord injury.
XX
PS   Example 5; Page 191-194; 270pp; English.
XX
CC   The invention relates to promoting locomotor recovery, inhibiting
CC   locomotor damage, limiting locomotor impairment, or limiting autonomic
CC   and sensory dysfunction following spinal cord injury by administering an
CC   anti-alpha_d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
CC   spinal cord injury victim. The method also involves the use of a ligand
CC   selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,
CC   vascular cell adhesion molecule). The method is useful for promoting
CC   locomotor recovery, inhibiting locomotor damage, limiting locomotor
CC   impairment, or limiting autonomic and sensory dysfunction following
CC   spinal cord injury. In particular, the spinal cord injury comprises
CC   compression of the spinal cord. The antibodies are also useful for
CC   reducing inflammation at the site of a central nervous system injury. The
CC   specification also details the identification of Beta2 integrin alphaD
CC   cDNAs and proteins, for use in raising the antibodies. Beta2 integrins
CC   are implicated in diseases such as LAD (leukocyte adhesion deficiency,
CC   inflammatory response, diabetes, multiple sclerosis, arthritis, graft
CC   atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative
CC   colitis, immune complex alveolitis and leukaemia. The present sequence is
CC   a Beta2 integrin alpha subunit sequence included for comparison with the
CC   Beta2 integrin alphaD protein sequences
XX
```

```
SQ   Sequence 1153 AA;

  Query Match             99.5%;  Score 5852;  DB 5;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:|||||||||||||:|||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              \\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 5
AAO14428
ID   AAO14428 standard; protein; 1153 AA.
XX
AC   AAO14428;
XX
DT   03-MAY-2002  (first entry)
XX
DE   Integrin Mac-1 alpha subunit.
XX
KW   Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
KW   open conformation; integrin related inflammatory disorder;
KW   integrin related immunological disorder; rheumatoid arthritis; ischaemia;
KW   reperfusion; hypovolemic shock; infarction; cerebral shock;
KW   viral infection; cancer; gene therapy; vaccine;
KW   bioactive agent screening.
XX
OS   Unidentified.
XX
PN   WO200204521-A2.
XX
PD   17-JAN-2002.
XX
PF   09-JUL-2001; 2001WO-US021805.
XX
PR   07-JUL-2000; 2000US-0216600P.
XX
PA   (CALY ) CALIFORNIA INST OF TECHNOLOGY.
PA   (BLOO-) CENT BLOOD RES.
XX
PI   Springer T;
XX
DR   WPI; 2002-148167/19.
XX
PT   New integrin I domain protein having alteration in at least 2
PT   noncontiguous regions and exits in an open conformation, useful for
PT   treating, preventing or suppressing inflammatory or immunological
PT   disorders.
XX
PS   Example 1; Fig 1F; 90pp; English.
XX
CC   The invention comprises structurally biased variant integrin inserted (I)
CC   domain proteins, wherein the alterations to the protein occur in at least
CC   two noncontinuous regions. Specifically the the variant integrin I domain
CC   proteins are structurally biased to exist in the open conformation,
CC   thereby altering the binding ability of the protein. The invention also
CC   comprises nucleic acids encoding the variant integrin I domain proteins.
CC   The integrin I domain proteins and nucleic acids are useful for treating,
CC   preventing or suppressing integrin related inflammatory and immunological
CC   disorders (e.g. rheumatoid arthritis). The variant integrin I domain
CC   proteins and nucleic acids can also be used for treating: ischaemia/
CC   reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral
CC   infection; and cancer. The variant integrin I domain nucleic acids and
CC   proteins may be used in gene therapy, as vaccines and to screen for
CC   bioactive agents. The present amino acid sequence represents the Mac-1
CC   alpha subunit of integrin
XX
SQ   Sequence 1153 AA;

  Query Match             99.5%;  Score 5852;  DB 5;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy        1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy       61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy      121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
            ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db      137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy      181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
            ||||||||||||||||||:||||||||||||||:||||||||||||||||||||||||| ::
Db      197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy      241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
            |||||||||||||||||| ||||||||||||||||||||||||:|||||||||||||:||
Db      257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy      301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy      361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy      421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy      481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy      541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy      601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy      661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy      721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy      781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy      841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy      901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy      961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy     1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy     1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX
AC ADD25615;
XX
DT 15-JAN-2004 (first entry)
XX
DE Binding domain-immunoglobulin fusion protein-associated protein #85.
XX
KW Binding domain; immunoglobulin; fusion protein; cytostatic;
KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;
KW neuroprotective; hinge region; immunoglobulin heavy chain;
KW CH2 constant region; CH3 constant region; IgG1;
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.
XX
OS Unidentified.
XX
PN US2003118592-A1.
XX
PD 26-JUN-2003.
XX
PF 25-JUL-2002; 2002US-00207655.
XX
PR 17-JAN-2001; 2001US-0367358P.
PR 17-JAN-2002; 2002US-00053530.
PR 03-JUN-2002; 2002US-0385691P.
XX
PA (GENE-) GENECRAFT INC.
XX
PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
XX
DR WPI; 2003-801317/75.
XX
PT New binding domain-immunoglobulin fusion protein, useful for treating a
PT subject having or suspected of having a malignant condition or a B-cell
PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX
PS Disclosure; SEQ ID NO 176; 157pp; English.
XX
CC The invention relates to a binding domain-immunoglobulin fusion protein
CC comprising a binding domain polypeptide that is fused to an
CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
CC CH2 constant region polypeptide that is fused to the hinge region
CC polypeptide, and an immunoglobulin heavy chain CH3 constant region
CC polypeptide that is fused to the CH2 constant region polypeptide. The
CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin
CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge
CC region polypeptide, derived from (a) having 3 or more cysteine residues;
CC where the mutated human IgG1 immunoglobulin hinge region polypeptide
CC contains 2 cysteine residues, where the first cysteine is not mutated; a
CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
CC (a) having 3 or more cysteine residues, where the mutated human IgG1
CC immunoglobulin hinge region polypeptide contains no more than one
CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
CC polypeptide, derived from (a) having 3 or more cysteine residues; where
CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains
CC no cysteine residues. The binding domain-immunoglobulin fusion protein is
CC capable of at least one immunological activity comprising antibody
CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The
CC binding domain polypeptide is capable of specifically binding to an
CC antigen. Also included are an isolated polynucleotide encoding the
CC binding domain-immunoglobulin fusion protein, a recombinant expression
CC construct comprising the polynucleotide (operably linked to a promoter),
CC a host cell transformed or transfected with a recombinant expression
CC construct, producing the binding domain-immunoglobulin fusion protein, a

```
CC   pharmaceutical composition comprising the binding domain-immunoglobulin
CC   fusion protein or polynucleotide and a carrier, and treating a subject
CC   having or suspected of having a malignant condition or a B-cell disorder.
CC   The binding domain-immunoglobulin fusion protein is useful for treating a
CC   subject having or suspected of having a malignant condition or a B-cell
CC   disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,
CC   myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple
CC   sclerosis or autoimmune disease. The present sequence is a binding domain
CC   -immunoglobulin fusion protein-associated protein sequence. Note: The
CC   sequence data for this patent formed part of the printed specification
CC   and is also available in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not
CC   identified the sequences in the printed specification by their SEQ ID
CC   number therefore none of the sequences can be explicitly identified.
XX
SQ   Sequence 1153 AA;

  Query Match             99.5%;  Score 5852;  DB 7;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:|||||||||||||||:|||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 7
AAR04136
ID   AAR04136 standard; protein; 1153 AA.
XX
AC   AAR04136;
XX
DT   09-SEP-2004  (revised)
DT   25-MAR-2003  (revised)
DT   07-SEP-1990  (first entry)
XX
DE   Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX
KW   Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
KW   non-specific defence system; integrin gene superfamily.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .16
FT                   /label= signal_peptide
FT   Modified-site   86. .88
FT                   /label= putative N-glycosylation site
FT   Modified-site   240. .242
FT                   /label= putative N-glycosylation site
FT   Modified-site   391. .393
FT                   /label= putative N-glycosylation site
FT   Modified-site   469. .471
FT                   /label= putative N-glycosylation site
FT   Modified-site   693. .695
FT                   /label= putative N-glycosylation site
FT   Modified-site   697. .699
FT                   /label= putative N-glycosylation site
FT   Modified-site   735. .737
FT                   /label= putative N-glycosylation site
FT   Modified-site   802. .804
FT                   /label= putative N-glycosylation site
FT   Modified-site   881. .883
FT                   /label= putative N-glycosylation site
FT   Modified-site   901. .903
FT                   /label= putative N-glycosylation site
FT   Modified-site   912. .914
FT                   /label= putative N-glycosylation site
FT   Modified-site   941. .943
FT                   /label= putative N-glycosylation site
FT   Modified-site   947. .949
FT                   /label= putative N-glycosylation site
FT   Modified-site   979. .981
FT                   /label= putative N-glycosylation site
```

```
FT   Modified-site   994. .996
FT                   /label= putative N-glycosylation site
FT   Modified-site   1022. .1024
FT                   /label= putative N-glycosylation site
FT   Modified-site   1045. .1047
FT                   /label= putative N-glycosylation site
FT   Modified-site   1051. .1053
FT                   /label= putative N-glycosylation site
FT   Modified-site   1076. .1078
FT                   /label= putative N-glycosylation site
FT   Region          1106. .1134
FT                   /label= putative_transmembrane_region
XX
PN   EP364690-A.
XX
PD   25-APR-1990.
XX
PF   17-AUG-1989;   89EP-00115159.
XX
PR   23-AUG-1988;   88US-00235353.
PR   09-MAR-1989;   89US-00321239.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Springer TA,  Corbi A;
XX
DR   WPI; 1990-125938/17.
DR   N-PSDB; AAQ04043.
XX
PT   New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating
PT   inflammation and viral infections, and in diagnosis.
XX
PS   Disclosure; Page ?; 3pp; English.
XX
CC   Mac-1 alpha subunit is involved in the response to inflammation, i.e.
CC   recognition of and migration to sites of inflammation. It also attaches
CC   to cellular substrates as part of this function making it useful in
CC   visualising endothelial tissue. Mac-1 is a member of the Integrin Gene
CC   superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25
CC   -MAR-2003 to correct PA field.)
CC
CC   Revised record issued on 09-SEP-2004 : Correction to feature table key
CC   and pages
XX
SQ   Sequence 1153 AA;

  Query Match             99.3%;  Score 5842;  DB 2;  Length 1153;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 1127;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:|||||||||||||:||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| |||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLLLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 8
ADM99589
ID   ADM99589 standard; protein; 1152 AA.
XX
AC   ADM99589;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Human integrin alphaM subunit precursor protein.
XX
KW   integrin alpha subunit; beta; antipsoriatic; thrombolytic; anticoagulant;
KW   osteopathic; cytostatic; immunosuppressive; antiinflammatory;
KW   neuroprotective; antisickling; immunotherapy; inflammatory;
KW   autoimmune disorder; thrombosis; cancer; osteoporosis;
KW   sickle cell anaemia; psoriasis; multiple sclerosis; human; precursor;
KW   alphaM.
XX
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   Misc-difference 965
FT                   /note= "Encoded by CCC"
XX
PN   WO2004007530-A2.
XX
PD   22-JAN-2004.
XX
PF   17-JUL-2003; 2003WO-US022301.
XX
PR   17-JUL-2002; 2002US-0396783P.
PR   17-JUL-2002; 2002US-0396790P.
PR   11-SEP-2002; 2002US-0410135P.
XX
PA   (BLOO-) CENT BLOOD RES INC.
XX
PI   Springer TA,  Takagi J;
XX
DR   WPI; 2004-122877/12.
DR   N-PSDB; ADM99588.
XX
PT   Novel modified integrin protein having extracellular domains of integrin
PT   alpha and beta subunits or integrin alphaII and beta3 subunit, useful for
PT   treating integrin mediated disorders.
XX
PS   Disclosure; SEQ ID NO 4; 232pp; English.
XX
CC   The invention relates to a novel isolated or recombinant modified
CC   integrin protein having extracellular domains of integrin alpha and beta
CC   subunits where one of the subunits has one or more mutations, an altered
CC   surface feature or an amino acid substitution or internal deletion,
CC   extracellular domains of the integrin beta subunit that comprise a
CC   mutation that alters a non-cysteine residue to cysteine or extracellular
CC   domains of integrin alpha and beta subunits. The polypeptide of the
CC   invention demonstrates antipsoriatic, thrombolytic, anticoagulant,
CC   osteopathic, cytostatic, immunosuppressive, antiinflammatory,
CC   neuroprotective and antisickling activities and may be useful for
CC   immunotherapy in order to prevent or treat an integrin-mediated disorder
CC   such as an inflammatory disorder, an autoimmune disorder, thrombosis,
CC   cancer, osteoporosis, sickle cell anaemia, psoriasis and multiple
CC   sclerosis. The current sequence is that of the human integrin alphaM
CC   subunit precursor protein of the invention.
XX
SQ   Sequence 1152 AA;

  Query Match             99.2%;  Score 5836.5;  DB 8;  Length 1152;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 1127;  Conservative    7;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:|||||||||||||:|||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| ||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 9
ADP12435
ID   ADP12435 standard; protein; 1152 AA.
XX
AC   ADP12435;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Protein encoded by mRNA of the invention #45.
XX
KW   transplant rejection; immune system; rheumatoid arthritis; lupus;
KW   inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX
OS   Homo sapiens.
XX
PN   WO2004042346-A2.
XX
PD   21-MAY-2004.
XX
PF   24-APR-2003; 2003WO-US012946.
```

XX
PR 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX
PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX
DR WPI; 2004-400724/37.
XX
PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
PS Claim 65; SEQ ID NO 2444; 1762pp; English.
XX
CC The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC protein that is encoded by the mRNA of the invention.
XX
SQ Sequence 1152 AA;

Query Match 99.1%; Score 5829.5; DB 8; Length 1152;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:||||||||||||:|||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| ||||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152
```

RESULT 10
ADP44061
ID ADP44061 standard; protein; 1163 AA.
XX
AC ADP44061;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human CD11C protein SEQ ID NO:14.
XX
KW HIV entry inhibitor; cell surface protein inhibitor; HIV infection;
KW anti-HIV; virucide; HIV uptake inhibitor; human; CD11C; integrin alpha X.
XX
OS Homo sapiens.
XX
PN WO2004053094-A2.
XX
PD 24-JUN-2004.
XX
PF 08-DEC-2003; 2003WO-US039208.
XX
PR 06-DEC-2002; 2002US-0431522P.
XX
PA (PPDP-) PPD DEV LP.
XX
PI Dunn SJ, Holzmayer TA;
XX
DR WPI; 2004-480928/45.
DR N-PSDB; ADP44060.
XX
PT Identifying an inhibitor of human immunodeficiency virus (HIV) entry into

PT a human host cell useful for preventing and/or treating HIV infection, by
PT identifying an inhibitor of a cell surface polypeptide such as CXCR-4.
XX
PS Claim 1; SEQ ID NO 14; 133pp; English.
XX
CC The present invention describes a method (M1) for identifying an
CC inhibitor of HIV entry into a human host cell. (M1) comprises identifying
CC an inhibitor of a cell surface polypeptide selected from CXCR-4 (352
CC amino acids, SEQ ID NO:8, ADP44055), CCR4 (360 amino acids, SEQ ID NO:10,
CC ADP44057), CCR7 (378 amino acids, SEQ ID NO:12, ADP44059), CD11C (1163
CC amino acids, SEQ ID NO:14, ADP44061), CD47 (323 amino acids, SEQ ID
CC NO:16, ADP44063), CD68 (354 amino acids, SEQ ID NO:18, ADP44065), CD69
CC (199 amino acids, SEQ ID NO:20, ADP44067), CD74 (566 amino acids, SEQ ID
CC NO:22, ADP44069), CSF3R (836 amino acids, SEQ ID NO:24, ADP44071), RARA
CC (462 amino acids, SEQ ID NO:26, ADP44073), GABBR1 (578 amino acids, SEQ
CC ID NO:28, ADP44075), P2X1 (150 amino acids, SEQ ID NO:30, ADP44077),
CC HELO1 (299 amino acids, SEQ ID NO:32, ADP44079), GPRK6 (576 amino acids,
CC SEQ ID NO:34, ADP44081), or PTK2B (1009 amino acids, SEQ ID NO:36,
CC ADP44083). Also described: (1) an inhibitor of a cell surface polypeptide
CC as described above in a human host cell preventing HIV entry into the
CC human host cell; (2) a pharmaceutical composition comprising an inhibitor
CC of (1) and a carrier; and (3) conferring resistance to HIV infection in
CC an individual, by administering the pharmaceutical composition of (2). An
CC inhibitor of HIV entry has anti-HIV and virucide activities, and can be
CC used as an HIV uptake inhibitor. (M1) is useful for identifying
CC protective compounds that inhibit entry of HIV into cells, useful for the
CC prevention and/or treatment of HIV infection. The present sequence
CC represents human integrin alpha X (CD11C), which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 1163 AA;

Query Match 59.0%; Score 3473; DB 8; Length 1163;
Best Local Similarity 61.0%; Pred. No. 1.4e-281;
Matches 689; Conservative 142; Mismatches 292; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||| |: :: ||| |||| | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||||||||||:|||| |||||||||| | | |: ||||| | | |
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || |||||||||   :| | :| |: | :: | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              : ||||:||: : || ||:  : || | |:||| :: ||  ||: : |||::| ||| ::
Db        198 QTHFTFEEFRRSSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAAKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:| || | |:|||| | |:||| |||| ||::  | :||| :|||| ::|:|::
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:|||||||||:| |||||| ||:|||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |  ||||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:|||||||||||||:|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDSDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||| ||||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
```

```
                ||  |  |||||||||||||:|| |||||||:|||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :||| :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :|||| ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:|| |||| ||| |||||| |:||||:| |:||| | |:||
Db        737 TLVGKPLLAFRNLRPMLAADAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |  |||| ||| ||   :::||  |||| || ||::||| |||:|||| |||:||| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| ||:||   ||||||: | |  | |  |:|||:||||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | ||| :|   :|: :| |  | ::: |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :     : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 11
ADN02004
ID ADN02004 standard; protein; 1163 AA.
XX
AC ADN02004;
XX
DT 17-JUN-2004 (first entry)
XX
DE Human inflammatory bowel disease marker - CD11c protein.
XX
KW marker gene; inflammatory bowel disease; IBD; FcgammaR IIIa;
KW FcgammaR IIIb; Mig; NRG-2; hexokinase; HM74; REG III; LPAP; Mip-1(beta);
KW L-selectin; EGFL6; IDO; IL-8; CD11c; TLR2; ulcerative colitis;
KW Crohn's disease.
XX
OS Homo sapiens.
XX
PN JP2004065120-A.
XX
PD 04-MAR-2004.
XX
PF 07-AUG-2002; 2002JP-00229705.
XX
PR 07-AUG-2002; 2002JP-00229705.
XX
PA (SUMU) SUMITOMO SEIYAKU KK.
XX
DR WPI; 2004-209124/20.
DR N-PSDB; ADN01990.
XX
PT Novel marker gene of inflammatory bowel disease (IBD) comprising base
PT sequence of e.g., Fc gamma R IIIa, Fc gamma R IIIb, Mig, NRG-2,
PT hexokinase 3, HM74, CD11c, TLR2 gene, useful for screening therapeutic

```
PT   agent for IBD.
XX
PS   Example; SEQ ID NO 27; 151pp; Japanese.
XX
CC   The invention comprises marker genes for inflammatory bowel disease (IBD)
CC   - FcgammaR IIIa, FcgammaR IIIb, Mig, NRG-2, hexokinase, HM74, REG III,
CC   LPAP, Mip-1(beta), L-selectin, EGFL6, IDO, IL-8, CD11c, and TLR2 genes.
CC   The DNA and protein sequences of the invention are useful in the
CC   detection and treatment of IBD (e.g. ulcerative colitis and Crohn's
CC   disease). The present amino acid sequence represents an IBD marker
CC   protein of the invention.
XX
SQ   Sequence 1163 AA;

  Query Match             58.8%;  Score 3459;  DB 8;  Length 1163;
  Best Local Similarity   60.9%;  Pred. No. 2.1e-280;
  Matches  687;  Conservative  142;  Mismatches  294;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||     |: :: ||| ||||    | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||||:|||| |||||||||| |  | |: ||||| |    |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || |||||||||   :|  |  :|  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              : ||||:||:   || ||:  : || | |:||| :: ||  ||: : |||::| ||| ::
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:| || | |:|||| |  |:||| |||| ||::  | :||| :|||| ::|:|::
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:|||||||||:| |||||| ||:|||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |   |||||||||||||| |    :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:||||||||||||::||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||| ||||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||   |  |||||||||||:|| |||||||:|||||||| |||||| |||:| ||||||
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :|||  :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:||  ||| |||  ||||||| |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | | | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::||  ||||  || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| ||:||   ||||||: | |  | |  |:|||:|||||  ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| |  | ::: |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :     : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:||||||||||| |||:|||||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143


RESULT 12
ADQ17510
ID   ADQ17510 standard; protein; 1163 AA.
XX
AC   ADQ17510;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated protein - SEQ ID 327.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 327; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   protein of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 1163 AA;
```

Query Match 58.8%; Score 3459; DB 8; Length 1163;
Best Local Similarity 60.9%; Pred. No. 2.1e-280;
Matches 687; Conservative 142; Mismatches 294; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 13
AAR07120
ID AAR07120 standard; protein; 1163 AA.
XX
AC AAR07120;
XX
DT 25-MAR-2003 (revised)
DT 05-FEB-1991 (first entry)
XX
DE p150.95 alpha subunit encoded by clone lambdaX47.
XX
KW p150,95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;
KW rhinovirus.
XX
OS Synthetic.
XX

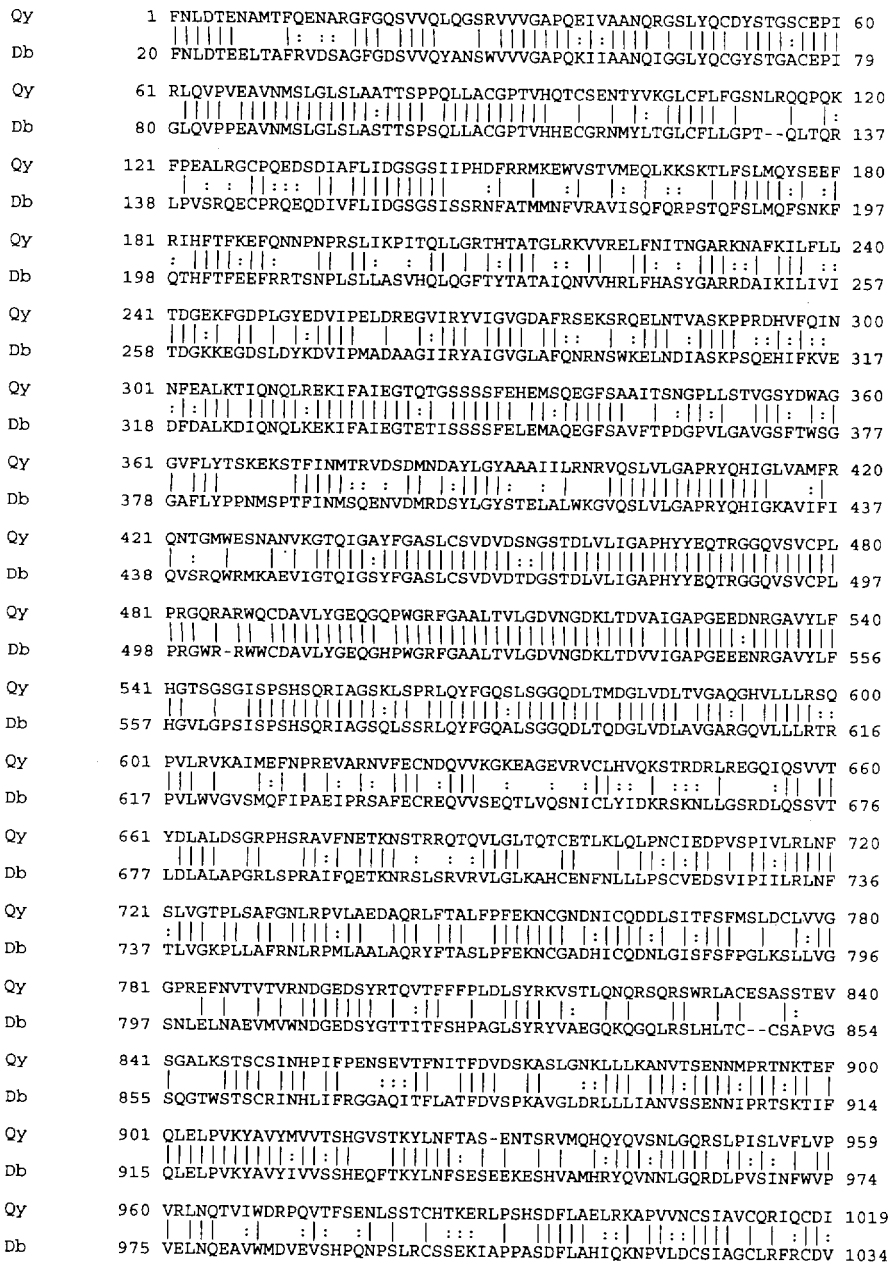
FH	Key	Location/Qualifiers
FT	Region	1. .19
FT		/label= signal peptide
FT	Region	20. .44
FT		/label= N-terminus
FT	Modified-site	61. .63
FT		/label= glycosylation site
FT	Modified-site	89. .91
FT		/label= glycosylation site
FT	Modified-site	385. .387
FT		/label= glycosylation site
FT	Modified-site	392. .394
FT		/label= glycosylation site
FT	Modified-site	697. .699
FT		/label= glycosylation site
FT	Modified-site	735. .737
FT		/label= glycosylation site
FT	Modified-site	899. .901
FT		/label= glycosylation site
FT	Modified-site	904. .906
FT		/label= glycosylation site
FT	Modified-site	939. .941
FT		/label= glycosylation site
FT	Modified-site	1050. .1052
FT		/label= glycosylation site
FT	Domain	1108. .1133
FT		/label= transmembrane

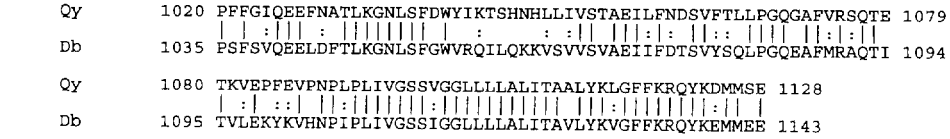
XX
PN WO9010646-A.
XX
PD 20-SEP-1990.
XX
PF 09-MAR-1990; 90WO-US001257.
XX
PR 09-MAR-1990; 90WO-US001257.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Corbi AA, Springer TA;
XX
DR WPI; 1990-304985/40.
DR N-PSDB; AAQ06068.
XX
PT Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-
PT unit of p150.95 cell surface adhesion receptor, opt. together with a beta
PT chain of CD-18 family.
XX
PS Disclosure; Fig 3; 59pp; English.
XX

```
CC   Clone lambda X47 was isolated from a cDNA library constructed from total
CC   RNA extracted from phorbol myristate acetate stimulated HL-60
CC   myelomonocytic cells. The library was screened with oligonucleotide
CC   probes based on tryptic peptide fragments of p150.95. The sequence can be
CC   attached to appropriate control elements and expressed in prokaryotic and
CC   eukaryotic cells. The protein can be used to treat or prevent rhinoviral
CC   infection because it interacts with ICAM-1 and inhibits cell-virus
CC   attachment. It can also be used as an anti-inflammatory agent. See also
CC   AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003
CC   to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1163 AA;

  Query Match             58.6%;  Score 3450;  DB 2;  Length 1163;
  Best Local Similarity   60.7%;  Pred. No. 1.2e-279;
  Matches  685;  Conservative  142;  Mismatches  296;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||    | |||||||:| |||| | |||| ||||:|||| 
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||||:|||| ||||||||||  |  | |: |||||  |    |  |: 
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || |||||||||   :|  |  :|  |: | ::  | |||||:| :| 
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              : ||||:||:   || ||:  : || | |:||| :: ||  ||: : |||::| ||| :: 
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:| || | |:|||| |  |:||| |||| ||::  | :||| :|||| ::|:|::  
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:|||||||||:| |||||| ||:|||||  | :||:|  |||: |:| 
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |   ||||||||||||| |   :| 
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:||||||||||||::|||||||||||||||||||||||||| 
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||||| |||||||:||||||| 
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||   |  |||||||||||||:|| |||||||:||||||| |||||  |||:| |||||:: 
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :|||  :   :  :||:: | ::: |    :|| || 
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| ||||| 
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:||  ||| |||  ||||||| |:||||:| |:|||   |  |:|| 
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |   ||  | |:|| 
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSA--PVA 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::||  ||||  || ||::||| |||:|||| |||:|| | 
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| ||:||   |||||: | |  | |  |:|||:||||| ||:|: | || 
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| | | ::: | ||||| ::| ||::|||| | | :||: 
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSQKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :     : :|| ||| |: ||:: ||||  |:|:|| 
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQERFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:| |||||||| |||:|||||||:|| | 
Db       1095 TVLEKYKVHNPTPLIVGSSIGALLLLALITAVLYKVGFFKRQYKEMMEE 1143


RESULT 14
AAW65091
ID   AAW65091 standard; protein; 1163 AA.
XX
AC   AAW65091;
XX
DT   28-SEP-1998  (first entry)
XX
DE   Human Beta-integrin CD11c subunit protein.
XX
KW   Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW   type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
KW   lung inflammation; acute respiratory distress syndrome; CD11c subunit;
KW   rheumatoid arthritis.
XX
OS   Homo sapiens.
XX
PN   US5728533-A.
XX
PD   17-MAR-1998.
XX
PF   07-JUN-1995;   95US-00485618.
XX
PR   23-DEC-1993;   93US-00173497.
PR   05-AUG-1994;   94US-00286889.
PR   21-DEC-1994;   94US-00362652.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Van Der Vieren M,  Gallatin WM;
XX
DR   WPI; 1998-206565/18.
XX
PT   Screening assay for modulators of integrin binding - using immobilised or
PT   labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.
XX
PS   Example 5; Fig 1A-D; 106pp; English.
XX
CC   This sequence represents a human beta-integrin CD11c subunit which is
CC   used to describe a method for identifying compounds that modulate the
CC   interaction of the beta-integrin alpha-d subunit with a binding partner
CC   of alpha-d which involves contacting an alpha-d polypeptide with an alpha
CC   -d binding partner, one of which is immobilised and the other of which is
CC   labelled, in the presence of a test compound, and determining if the
CC   compound affects binding between the alpha-d polypeptide and alpha-d
CC   binding partner, where the alpha-d polypeptide is alpha-d or its fragment
CC   comprising the cytoplasmic, transmembrane or extracellular domain of
CC   alpha-d. Compounds that modulate alpha-d binding could be used to treat
CC   diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC   asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC   and rheumatoid arthritis
XX
```

SQ Sequence 1163 AA;

Query Match 58.4%; Score 3436; DB 2; Length 1163;
Best Local Similarity 60.6%; Pred. No. 1.8e-278;
Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;





RESULT 15
AAB07361
ID AAB07361 standard; protein; 1163 AA.
XX
AC AAB07361;
XX
DT 17-JAN-2001 (first entry)
XX
DE Human CD11c protein sequence.
XX
KW Human; macrophage infiltration inhibition; alpha_d integrin;
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW rheumatoid arthritis; central nervous system injury; CD11c.
XX
OS Homo sapiens.
XX
PN WO200029446-A1.
XX
PD 25-MAY-2000.
XX
PF 16-NOV-1999; 99WO-US027139.
XX
PR 16-NOV-1998; 98US-00193043.
PR 08-JUL-1999; 99US-00350259.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gallatin MW, Van Der Vieren M;
XX
DR WPI; 2000-387751/33.
XX
PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous system
PT injury sites.
XX
PS Example 5; Fig 1; 270pp; English.
XX
CC Integrins are a class of membrane-associated molecules that participate
CC in cellular adhesion. Integrins are made up of an alpha subunit and a
CC beta subunit. One class of human integrins are restricted to expression
CC in white blood cells and have a common beta2 subunit: the leukocyte
CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC have an important role in immune and inflammatory responses. The present
CC protein sequence is the human integrin alpha subunit CD11c. This sequence
CC was used in an alignment to identify a novel beta2 integrin alpha
CC subunit: alpha_d (AAA60014 and AAB07359). The present sequence has
CC approximately 66% identity to the protein sequence of alpha_d. The
CC Alpha_d gene and protein may be useful in therapy for diseases linked to
CC aberrant alpha_d function e.g. Type I diabetes, atherosclerosis, multiple
CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
CC (LAD). In addition, anti-alpha_d monoclonal antibodies may be used in the
CC inhibition of macrophage infiltration at the site of a central nervous
CC system injury. The monoclonal antibodies can also be used to detect and
CC diagnose Crohn's disease
XX
SQ Sequence 1163 AA;

Query Match 58.4%; Score 3436; DB 3; Length 1163;
Best Local Similarity 60.6%; Pred. No. 1.8e-278;

Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAIKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        318 DFDALKDIQNQLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        438 QVSRQWRMKAEVIGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALAPGRLSPRAIFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTC--CSAPVG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYIVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEIIFDTSVYSQLPGQEAFMRAQTI 1094

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPIPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

Search completed: November 9, 2004, 12:08:19
Job time : 108.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 12:17:49 ; Search time 97.5 Seconds
(without alignments)
4119.157 Million cell updates/sec

Title: US-09-902-481B-4
Perfect score: 5884
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5884	100.0	1137	10	US-09-902-481A-4	Sequence 4, Appli
2	5857	99.5	1137	10	US-09-902-481A-3	Sequence 3, Appli
3	5852	99.5	1153	9	US-09-350-259-3	Sequence 3, Appli
4	5852	99.5	1153	10	US-09-902-481A-1	Sequence 1, Appli
5	5852	99.5	1153	10	US-09-891-943-3	Sequence 3, Appli
6	5852	99.5	1153	14	US-10-144-259-30	Sequence 30, Appl
7	5852	99.5	1153	14	US-10-207-655-176	Sequence 176, App
8	5851	99.4	1137	10	US-09-902-481A-5	Sequence 5, Appli
9	5845	99.3	1137	10	US-09-902-481A-6	Sequence 6, Appli
10	5836.5	99.2	1152	9	US-09-945-265-4	Sequence 4, Appli
11	3459	58.8	1163	14	US-10-116-275-204	Sequence 204, App
12	3436	58.4	1163	9	US-09-350-259-4	Sequence 4, Appli
13	3436	58.4	1163	10	US-09-891-943-4	Sequence 4, Appli
14	3401	57.8	1161	9	US-09-350-259-2	Sequence 2, Appli
15	3401	57.8	1161	10	US-09-891-943-2	Sequence 2, Appli
16	3385.5	57.5	1161	9	US-09-350-259-99	Sequence 99, Appl
17	3385.5	57.5	1161	10	US-09-891-943-99	Sequence 99, Appl
18	3224.5	54.8	1161	9	US-09-350-259-55	Sequence 55, Appl
19	3224.5	54.8	1161	10	US-09-891-943-55	Sequence 55, Appl
20	3212	54.6	1161	9	US-09-350-259-53	Sequence 53, Appl
21	3212	54.6	1161	10	US-09-891-943-53	Sequence 53, Appl
22	3203.5	54.4	1151	9	US-09-350-259-37	Sequence 37, Appl
23	3203.5	54.4	1151	10	US-09-891-943-37	Sequence 37, Appl
24	3191	54.2	1155	9	US-09-350-259-46	Sequence 46, Appl
25	3191	54.2	1155	10	US-09-891-943-46	Sequence 46, Appl
26	1848	31.4	369	13	US-10-087-192-1212	Sequence 1212, Ap
27	1547.5	26.3	1170	9	US-09-945-265-2	Sequence 2, Appli
28	1543.5	26.2	1170	15	US-10-261-164-1	Sequence 1, Appli
29	1520	25.8	1223	16	US-10-408-765A-295	Sequence 295, App
30	1350.5	23.0	1086	16	US-10-408-765A-1871	Sequence 1871, Ap
31	1229.5	20.9	494	9	US-09-350-259-103	Sequence 103, App
32	1229.5	20.9	494	10	US-09-891-943-103	Sequence 103, App
33	1141.5	19.4	413	9	US-09-350-259-101	Sequence 101, App
34	1141.5	19.4	413	10	US-09-891-943-101	Sequence 101, App
35	1140	19.4	1179	14	US-10-177-550-2	Sequence 2, Appli
36	1140	19.4	1179	14	US-10-173-551-2	Sequence 2, Appli
37	1093.5	18.6	1151	10	US-09-984-130-103	Sequence 103, App
38	1093.5	18.6	1151	10	US-09-836-353A-103	Sequence 103, App
39	1093.5	18.6	1179	10	US-09-918-715-250	Sequence 250, App
40	1093.5	18.6	1179	17	US-10-474-794-250	Sequence 250, App
41	1092	18.6	1188	14	US-10-291-265-810	Sequence 810, App
42	1088	18.5	1188	14	US-10-291-265-338	Sequence 338, App
43	1084.5	18.4	1189	10	US-09-984-130-35	Sequence 35, Appl
44	1084.5	18.4	1189	10	US-09-836-353A-35	Sequence 35, Appl
45	1084.5	18.4	1189	15	US-10-262-839-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-902-481A-4
; Sequence 4, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4

Query Match 100.0%; Score 5884; DB 10; Length 1137;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
```

```
Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

RESULT 2
US-09-902-481A-3
; Sequence 3, Application US/09902481A

; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 99.5%; Score 5857; DB 10; Length 1137;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1131; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||| |||:|||||||||||||||||||||||
Db        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||| |
Db        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
```

```
Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137


RESULT 3
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT: Van der Vieren, Monica
;  TITLE OF INVENTION: No. US20020062008A1el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/350,259
;  CURRENT FILING DATE: 1999-07-08
;  EARLIER APPLICATION NUMBER: 09/193,043
;  EARLIER FILING DATE: 1998-11-16
;  EARLIER APPLICATION NUMBER: 08/173,497
;  EARLIER FILING DATE: 1993-12-23
;  EARLIER APPLICATION NUMBER: 08/286,889
;  EARLIER FILING DATE: 1994-08-05
;  EARLIER APPLICATION NUMBER: 08/362,652
;  EARLIER FILING DATE: 1994-12-21
;  EARLIER APPLICATION NUMBER: 08/943,363
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-350-259-3

  Query Match           99.5%;  Score 5852;  DB 9;  Length 1153;
  Best Local Similarity  99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||

Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| ||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 4
US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
;  APPLICANT: Springer, Timothy
```

```
;  APPLICANT:  Shimaoka, Motomu
;  APPLICANT:  Shifman, Julia
;  APPLICANT:  Mayo, Stephen
;  TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
;  FILE REFERENCE: A-70586-1/RFT/RMS/RMK
;  CURRENT APPLICATION NUMBER: US/09/902,481A
;  CURRENT FILING DATE:  2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/216,600
;  PRIOR FILING DATE: 2000-07-07
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (17)..()
;   OTHER INFORMATION:
US-09-902-481A-1

  Query Match             99.5%;  Score 5852;  DB 10;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||:||||||||||||||:||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| ||||||||||||||||||||||||:||||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 5
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT:  Van der Vieren, Monica
;  TITLE OF INVENTION: No. US20030077278A1el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/891,943
;  CURRENT FILING DATE:  2001-06-26
;  PRIOR APPLICATION NUMBER: 09/193,043
;  PRIOR FILING DATE: 1998-11-16
;  PRIOR APPLICATION NUMBER: 08/286,889
;  PRIOR FILING DATE: 1994-08-05
;  PRIOR APPLICATION NUMBER: 08/362,652
;  PRIOR FILING DATE: 1994-12-21
;  PRIOR APPLICATION NUMBER: 08/943,363
;  PRIOR FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-891-943-3

  Query Match             99.5%;  Score 5852;  DB 10;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
```

```
                    |||||||||||||||||:|||||||||||||||:|||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 6
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
;  APPLICANT: Arnaout, M. Amin
;  APPLICANT:  Li, Rui
;  APPLICANT:  Xiong, Jian-Ping
;  TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: 00786-548001
;  CURRENT APPLICATION NUMBER: US/10/144,259
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: US 09/758,493
;  PRIOR FILING DATE: 2001-01-11
;  PRIOR APPLICATION NUMBER: US 60/221,950
;  PRIOR FILING DATE: 2000-07-31
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-144-259-30

  Query Match             99.5%;  Score 5852;  DB 14;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:|||||||||||||:|||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 7
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 99.5%; Score 5852; DB 14; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180

Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 8
US-09-902-481A-5
; Sequence 5, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic

US-09-902-481A-5

Query Match 99.4%; Score 5851; DB 10; Length 1137;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||||| ||::|||||||||||||||||||||||
Db        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||| |:
Db        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
```

```
Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

RESULT 9
US-09-902-481A-6
; Sequence 6, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6

Query Match 99.3%; Score 5845; DB 10; Length 1137;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:|||||||||||||:|||:||| |||||||||||||| ::
Db        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||||||||:|||||||||||:|
Db        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137


RESULT 10
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
;  APPLICANT: Springer, Timothy A.
;  APPLICANT:  Shimaoka, Motomu
;  APPLICANT:  Lu, Chafen
;  TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
;  TITLE OF INVENTION:  DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
;  FILE REFERENCE: CBN-002CP
;  CURRENT APPLICATION NUMBER: US/09/945,265
;  CURRENT FILING DATE:  2001-08-31
;  PRIOR APPLICATION NUMBER: US 60/229,700
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 1152
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-945-265-4

  Query Match            99.2%;  Score 5836.5;  DB 9;  Length 1152;
  Best Local Similarity  99.1%;  Pred. No. 0;
  Matches 1127;  Conservative    7;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180

Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152
```

```
RESULT 11
US-10-116-275-204
; Sequence 204, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
;  APPLICANT: Elan Pharmaceutical Technology
;  APPLICANT:  O'Mahony, Daniel J.
;  APPLICANT:  Brayden, David
;  APPLICANT:  Byrne, Daragh
;  APPLICANT:  Lambkin, Imelda
;  APPLICANT:  Higgins, Lisa
;  TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
;  TITLE OF INVENTION:  Compositions Targeting Peyer's Patches and M Cell Receptors
;  FILE REFERENCE: E1067/20087
;  CURRENT APPLICATION NUMBER: US/10/116,275
;  CURRENT FILING DATE:  2002-10-04
;  NUMBER OF SEQ ID NOS: 349
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 204
;   LENGTH: 1163
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-116-275-204

  Query Match             58.8%;  Score 3459;  DB 14;  Length 1163;
  Best Local Similarity   60.9%;  Pred. No. 6.3e-286;
  Matches  687;  Conservative  142;  Mismatches  294;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||   | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||||||||||:|||| ||||||||||   |  | |: ||||| |    | |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || |||||||||   :|  |  :|  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              : ||||:||:   || ||:  : || | |:||| :: ||  ||: : |||::| ||| ::
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:| || | |:|||   |  |:||| |||| ||::  | :||| :|||| ::|:|::
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:|||||||:| |||||| ||:||||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |   |||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |    | | |||||:|||||||||||::|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  ||||||||||||:|| |||||||:||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :|||  :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECRERQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:||  ||| ||| |||||||  |:||||:|  |:||| |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::|| ||||  || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| ||:||  ||||||: | |  | |  |:|||:|||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| |  | :::  |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :      : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143


RESULT 12
US-09-350-259-4
; Sequence 4, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT: Van der Vieren, Monica
;  TITLE OF INVENTION: No. US20020062008A1el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/350,259
;  CURRENT FILING DATE: 1999-07-08
;  EARLIER APPLICATION NUMBER: 09/193,043
;  EARLIER FILING DATE: 1998-11-16
;  EARLIER APPLICATION NUMBER: 08/173,497
;  EARLIER FILING DATE: 1993-12-23
;  EARLIER APPLICATION NUMBER: 08/286,889
;  EARLIER FILING DATE: 1994-08-05
;  EARLIER APPLICATION NUMBER: 08/362,652
;  EARLIER FILING DATE: 1994-12-21
;  EARLIER APPLICATION NUMBER: 08/943,363
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 1163
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-350-259-4

  Query Match             58.4%;  Score 3436;  DB 9;  Length 1163;
  Best Local Similarity   60.6%;  Pred. No. 5.9e-284;
  Matches  684;  Conservative  148;  Mismatches  291;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||   | |||||||:|:|||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||||||||||:|||| ||||||||||   |  | |: ||||| |    | |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137
```

```
Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAIKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        318 DFDALKDIQNQLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        438 QVSRQWRMKAEVIGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALAPGRLSPRAIFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTC--CSAPVG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYIVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEIIFDTSVYSQLPGQEAFMRAQTI 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPIPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143

RESULT 13
US-09-891-943-4
; Sequence 4, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT:  Van der Vieren, Monica
;  TITLE OF INVENTION: No. US20030077278A1el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/891,943
;  CURRENT FILING DATE:  2001-06-26
;  PRIOR APPLICATION NUMBER: 09/193,043
;  PRIOR FILING DATE: 1998-11-16
;  PRIOR APPLICATION NUMBER: 08/286,889
;  PRIOR FILING DATE: 1994-08-05
;  PRIOR APPLICATION NUMBER: 08/362,652
;  PRIOR FILING DATE: 1994-12-21
;  PRIOR APPLICATION NUMBER: 08/943,363
;  PRIOR FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 1163
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-891-943-4

  Query Match           58.4%;  Score 3436;  DB 10;  Length 1163;
  Best Local Similarity  60.6%;  Pred. No. 5.9e-284;
  Matches  684;  Conservative  148;  Mismatches  291;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAIKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        318 DFDALKDIQNQLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        438 QVSRQWRMKAEVIGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
```

```
Db        677 LDLALAPGRLSPRAIFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTC--CSAPVG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959

Db        915 QLELPVKYAVYIVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019

Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079

Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEIIFDTSVYSQLPGQEAFMRAQTI 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128

Db       1095 TVLEKYKVHNPIPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 14
US-09-350-259-2
; Sequence 2, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-2

Query Match 57.8%; Score 3401; DB 9; Length 1161;
Best Local Similarity 59.3%; Pred. No. 5.8e-281;
Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQTGRLYDCAAATGMCQPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-EIIQT 135

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180

Db        136 VPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFEGTDTLFALMQYSNLL 195

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Db        196 KIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHHKNGARKSAKKILIVI 255

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Db        256 TDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELNTISSAPPQDHVFKVD 315

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        316 NFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGSFSWSG 375

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        376 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVLGAPRYQHTGKAVIFT 435

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        436 QVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPHYYEQTRGGQVSVCPL 495

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        496 PRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAIGAPGEQENRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLLRSL 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        616 PVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        674 FDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCVEDVVSPIILHLNF 733

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        734 SLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDLGVTLSFSGLQTLTVG 793

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        794 SSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        853 EG-LRSSRCSVNHPIFHEGSNGTFIVTFDVSYKATLGDRMLMRASASSENNKASSSKATF 911

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNF-TASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959

Db        912 QLELPVKYAVYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQRDLAISINFWVP 971

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019

Db        972 VLLNGVAVWDVVMEAPSQSL--PCVSERKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1029

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079

Db       1030 PSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQLPGQEAFMRAQME 1089

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128

Db       1090 MVLEEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRHYKEMLED 1138
```

RESULT 15
US-09-891-943-2
; Sequence 2, Application US/09891943
; Publication No. US20030077278A1

```
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT:  Van der Vieren, Monica
;  TITLE OF INVENTION: No. US20030077278A1el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/891,943
;  CURRENT FILING DATE:  2001-06-26
;  PRIOR APPLICATION NUMBER: 09/193,043
;  PRIOR FILING DATE: 1998-11-16
;  PRIOR APPLICATION NUMBER: 08/286,889
;  PRIOR FILING DATE: 1994-08-05
;  PRIOR APPLICATION NUMBER: 08/362,652
;  PRIOR FILING DATE: 1994-12-21
;  PRIOR APPLICATION NUMBER: 08/943,363
;  PRIOR FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1161
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-891-943-2

  Query Match             57.8%;  Score 3401;  DB 10;  Length 1161;
  Best Local Similarity   59.3%;  Pred. No. 5.8e-281;
  Matches  669;  Conservative  167;  Mismatches  285;  Indels    8;  Gaps    6;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    |||:| |||||||  |||:|||| |:|||| | || |  :|| |:||
Db         17 FNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQTGRLYDCAAATGMCQPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               | :  |||||||||:|||:|:  :|||||||:|: | ||:| || | | ||    : |
Db         77 PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-EIIQT 135

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              |:|   || :: || |||||||||  :|| :|| :|  || | : : |||:|||||
Db        136 VPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFEGTDTLFALMQYSNLL 195

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              :||||| :|: :|: :||: || || | | ||||:  || :||:  |||||:| ||| ::
Db        196 KIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHHKNGARKSAKKILIVI 255

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:|: ||| | ||||: :: |:||| |||| ||:   :||||||::| ||:||||:::
Db        256 TDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELNTISSAPPQDHVFKVD 315

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              || || :|| ||:|||:|:||||: :||||:||||||||| |:| :|  |  |||: |:|
Db        316 NFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGSFSWSG 375

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |   ||:|||||||||| |   :|
Db        376 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVLGAPRYQHTGKAVIFT 435

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:|||||||||||||:|||||:|||||||||||||||||||||
Db        436 QVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPHYYEQTRGGQVSVCPL 495

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||| :|||||| |||| ||||||||||||||| ||| |||||||||::||||||||
Db        496 PRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAIGAPGEQENRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || | |||||||||||| |:|||||||||:|||||||| |||:|| |||:| ||||||
Db        556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLLRSL 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |||:|    | |:| |||: |: | ::     |||:  ||| :|||: |:|  | ||| |
Db        616 PVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              :||||| ||  |||:|||||| |  : : |||   |||||| ||:|:|| ||||:| |||

Db        674 FDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCVEDVVSPIILHLNF 733

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||  |: :  |||||||  :| ||||  |||||||  | :|: || :| ||  |  | ||
Db        734 SLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDLGVTLSFSGLQTLTVG 793

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
                | || ||| | ||||| | |: ::|  ||:|:||  | |  | : |||||:   ||
Db        794 SSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              | |:|: ||:|||||  | |  || :||||  ||:||:::|::|: :||||   ::|  |
Db        853 EG-LRSSRCSVNHPIFHEGSNGTFIVTFDVSYKATLGDRMLMRASASSENNKASSSKATF 911

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNF-TASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| :::   |||| || |: |   :  :|:|:|:|:|| || | ||: | ||
Db        912 QLELPVKYAVYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQRDLAISINFWVP 971

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | ||   :||      |::|   | :: : | ||||| :: ::|:::|||| | : :||:
Db        972 VLLNGVAVWDVVMEAPSQSL--PCVSERKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1029

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| | :|    :|:|| ||| |: ||:: |||| ||:|:| |
Db       1030 PSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQLPGQEAFMRAQME 1089

Qy       1080 TKVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
                :|  || | :|:|:|||||  |||||||||  |||||||||  ||:|: :
Db       1090 MVLEEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRHYKEMLED 1138


Search completed: November  9, 2004, 12:46:53
Job time : 102.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:57 ; Search time 23 Seconds
(without alignments)
3278.416 Million cell updates/sec

Title: US-09-902-481B-4
Perfect score: 5884
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5852	99.5	1153	1	US-08-173-497-3	Sequence 3, Appli
2	5852	99.5	1153	1	US-08-286-889-3	Sequence 3, Appli
3	5852	99.5	1153	1	US-08-485-618-3	Sequence 3, Appli
4	5852	99.5	1153	1	US-08-362-652-3	Sequence 3, Appli
5	5852	99.5	1153	2	US-08-605-672-3	Sequence 3, Appli
6	5852	99.5	1153	2	US-08-482-293A-3	Sequence 3, Appli
7	5852	99.5	1153	2	US-08-943-363-3	Sequence 3, Appli
8	5852	99.5	1153	3	US-09-193-043-3	Sequence 3, Appli
9	5852	99.5	1153	4	US-09-688-307A-3	Sequence 3, Appli
10	5852	99.5	1153	4	US-09-350-259-3	Sequence 3, Appli
11	5821.5	98.9	1152	2	US-08-476-062A-43	Sequence 43, Appl
12	5821.5	98.9	1152	5	PCT-US96-01314-43	Sequence 43, Appl
13	5821.5	98.9	1152	6	5424399-2	Patent No. 5424399
14	3459	58.8	1163	2	US-08-476-062A-44	Sequence 44, Appl
15	3459	58.8	1163	5	PCT-US96-01314-44	Sequence 44, Appl
16	3436	58.4	1163	1	US-08-173-497-4	Sequence 4, Appli
17	3436	58.4	1163	1	US-08-286-889-4	Sequence 4, Appli
18	3436	58.4	1163	1	US-08-485-618-4	Sequence 4, Appli
19	3436	58.4	1163	1	US-08-362-652-4	Sequence 4, Appli
20	3436	58.4	1163	2	US-08-605-672-4	Sequence 4, Appli
21	3436	58.4	1163	2	US-08-482-293A-4	Sequence 4, Appli
22	3436	58.4	1163	2	US-08-943-363-4	Sequence 4, Appli
23	3436	58.4	1163	3	US-09-193-043-4	Sequence 4, Appli
24	3436	58.4	1163	4	US-09-688-307A-4	Sequence 4, Appli
25	3436	58.4	1163	4	US-09-350-259-4	Sequence 4, Appli
26	3401	57.8	1161	1	US-08-173-497-2	Sequence 2, Appli
27	3401	57.8	1161	1	US-08-286-889-2	Sequence 2, Appli
28	3401	57.8	1161	1	US-08-485-618-2	Sequence 2, Appli
29	3401	57.8	1161	1	US-08-362-652-2	Sequence 2, Appli
30	3401	57.8	1161	2	US-08-605-672-2	Sequence 2, Appli
31	3401	57.8	1161	2	US-08-482-293A-2	Sequence 2, Appli
32	3401	57.8	1161	2	US-08-943-363-2	Sequence 2, Appli
33	3401	57.8	1161	3	US-09-193-043-2	Sequence 2, Appli
34	3401	57.8	1161	4	US-09-688-307A-2	Sequence 2, Appli
35	3401	57.8	1161	4	US-09-350-259-2	Sequence 2, Appli
36	3385.5	57.5	1161	1	US-08-485-618-99	Sequence 99, Appl
37	3385.5	57.5	1161	2	US-08-605-672-99	Sequence 99, Appl
38	3385.5	57.5	1161	2	US-08-482-293A-99	Sequence 99, Appl
39	3385.5	57.5	1161	2	US-08-943-363-99	Sequence 99, Appl
40	3385.5	57.5	1161	3	US-09-193-043-99	Sequence 99, Appl
41	3385.5	57.5	1161	4	US-09-688-307A-99	Sequence 99, Appl
42	3385.5	57.5	1161	4	US-09-350-259-99	Sequence 99, Appl
43	3224.5	54.8	1161	3	US-09-193-043-55	Sequence 55, Appl
44	3224.5	54.8	1161	4	US-09-688-307A-55	Sequence 55, Appl
45	3224.5	54.8	1161	4	US-09-350-259-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76
```

```
Qy        61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy       121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db       137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy       181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
             |||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db       197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy       241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
             |||||||||||||||| |||||||||||||||||||||||:||||||||||:|:|
Db       257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy       301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy       361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy       421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy       481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy       541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy       601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy       661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy       721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy       781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy       841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy       901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy       961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy      1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy      1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1097 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Mich
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5470953el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,889
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  P38,659
;      REFERENCE/DOCKET NUMBER:  27866/32168
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-286-889-3

  Query Match            99.5%;  Score 5852;  DB 1;  Length 1153;
  Best Local Similarity  99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| |||||||||||||||||||||||:||||||||||:|:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376
```

```
Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 PFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 PFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 3
US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5728533el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,618
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32797
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-485-618-3

  Query Match             99.5%;  Score 5852;  DB 1;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| |||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 4
US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-362-652-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| ||||||||||||||||||||||:||||||||||:|||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 5
US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5817515el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/605,672
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32684
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-605-672-3
```

```
Query Match             99.5%;  Score 5852;  DB 2;  Length 1153;
Best Local Similarity   99.2%;  Pred. No. 0;
Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036
```

```
Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 6
US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5831029el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,293A
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32684
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-482-293A-3

  Query Match             99.5%;  Score 5852;  DB 2;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||||:|||||||||||||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| ||||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5837478el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/943,363
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32684
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-943-363-3

  Query Match             99.5%;  Score 5852;  DB 2;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153

RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT: Van der Vieren, Monica
;  TITLE OF INVENTION: No. 6251395el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/193,043
;  CURRENT FILING DATE: 1998-11-16
;  EARLIER APPLICATION NUMBER: 08/173,497
;  EARLIER FILING DATE: 1993-12-23
;  EARLIER APPLICATION NUMBER: 08/286,889
;  EARLIER FILING DATE: 1994-08-05
;  EARLIER APPLICATION NUMBER: 08/362,652
;  EARLIER FILING DATE: 1994-12-21
;  EARLIER APPLICATION NUMBER: 08/943,363
;  EARLIER FILING DATE: 1997-10-03
```

```
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-193-043-3

  Query Match             99.5%;  Score 5852;  DB 3;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||||||:||||||||||:|||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 9
US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT:  Van der Vieren, Monica
;  TITLE OF INVENTION: No. 6432404el Human Beta-2
;  FILE REFERENCE: 27866/36646
;  CURRENT APPLICATION NUMBER: US/09/688,307A
;  CURRENT FILING DATE:  2000-10-13
;  PRIOR APPLICATION NUMBER: 09/193,043
;  PRIOR FILING DATE: 1998-11-16
;  PRIOR APPLICATION NUMBER: 08/605,672
;  PRIOR FILING DATE: 1996-02-22
;  PRIOR APPLICATION NUMBER: 08/173,497
;  PRIOR FILING DATE: 1993-12-23
;  PRIOR APPLICATION NUMBER: 08/286,889
;  PRIOR FILING DATE: 1994-08-05
;  PRIOR APPLICATION NUMBER: 08/362,652
;  PRIOR FILING DATE: 1994-12-21
;  PRIOR APPLICATION NUMBER: 08/943,363
;  PRIOR FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-688-307A-3

  Query Match             99.5%;  Score 5852;  DB 4;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||||||:||||||||||:|||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 10
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT: Van der Vieren, Monica
;  TITLE OF INVENTION: No. 6620915el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/350,259
;  CURRENT FILING DATE: 1999-07-08
;  EARLIER APPLICATION NUMBER: 09/193,043
;  EARLIER FILING DATE: 1998-11-16
;  EARLIER APPLICATION NUMBER: 08/173,497
;  EARLIER FILING DATE: 1993-12-23
;  EARLIER APPLICATION NUMBER: 08/286,889
;  EARLIER FILING DATE: 1994-08-05
;  EARLIER APPLICATION NUMBER: 08/362,652
;  EARLIER FILING DATE: 1994-12-21
;  EARLIER APPLICATION NUMBER: 08/943,363
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-350-259-3

  Query Match             99.5%;  Score 5852;  DB 4;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              |||||||||||||||||||:|||||||||||:|||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 11
US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
;  GENERAL INFORMATION:
;    APPLICANT:  Arnaout, M. Amin
;    TITLE OF INVENTION:  CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
;    TITLE OF INVENTION:  RESPONSES WITH BETA2 INTEGRINS
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,062A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/216,081
;      FILING DATE:  21-MAR-1994
;      APPLICATION NUMBER:  07/637,830
;      FILING DATE:  04-JAN-1991
;      APPLICATION NUMBER:  07/539,842
;      FILING DATE:  18-JUN-1990
;      APPLICATION NUMBER:  07/212,573
;      FILING DATE:  28-JUN-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Freeman, John W.
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  00786/068003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1152 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-476-062A-43

  Query Match             98.9%;  Score 5821.5;  DB 2;  Length 1152;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1125;  Conservative    7;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180

Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        796 GPRESNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 12
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
;  GENERAL INFORMATION:
;    APPLICANT:  M. Amin Arnaout
;    TITLE OF INVENTION:  METHODS FOR IDENTIFYING INTEGRIN
;    TITLE OF INVENTION:  ANTAGONISTS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/01314
;      FILING DATE:  30-JAN-96
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/380,167
;      FILING DATE:  30-JAN-95
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  John W. Freeman
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  00786/267001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1152
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
PCT-US96-01314-43

  Query Match             98.9%;  Score 5821.5;  DB 5;  Length 1152;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1125;  Conservative    7;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:||||||||||||||:||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| ||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPRESNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 13
5424399-2
;Patent No. 5424399
;    APPLICANT: ARNAOUT, M. AMIN
;    TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
;    NUMBER OF SEQUENCES: 12
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/78,871
;      FILING DATE: 16-JUN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 539,842
;      FILING DATE: 18-JUN-1990
;      APPLICATION NUMBER: 212,573
;      FILING DATE: 28-JUN-1988
;SEQ ID NO:2:
;      LENGTH: 1152
5424399-2

  Query Match             98.9%;  Score 5821.5;  DB 6;  Length 1152;
```

```
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1125;  Conservative    7;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
              ||||||||||||||||||:||||||||||||:||||||||||||||||||||||||| ::
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||| ||||||||||||||||||||||:||||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPRESNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 14
US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
;  GENERAL INFORMATION:
;    APPLICANT:  Arnaout, M. Amin
;    TITLE OF INVENTION:  CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
;    TITLE OF INVENTION:  RESPONSES WITH BETA2 INTEGRINS
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,062A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/216,081
;      FILING DATE:  21-MAR-1994
;      APPLICATION NUMBER:  07/637,830
;      FILING DATE:  04-JAN-1991
;      APPLICATION NUMBER:  07/539,842
;      FILING DATE:  18-JUN-1990
;      APPLICATION NUMBER:  07/212,573
;      FILING DATE:  28-JUN-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Freeman, John W.
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  00786/068003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1163 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-476-062A-44

  Query Match             58.8%;  Score 3459;  DB 2;  Length 1163;
  Best Local Similarity   60.9%;  Pred. No. 1.2e-289;
  Matches  687;  Conservative  142;  Mismatches  294;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||   | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||||||||||:|||| ||||||||||   |  | |: ||||| |     |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || |||||||||   :|  |  :|  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240
```

```
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959

Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019

Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079

Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128

Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143


RESULT 15
PCT-US96-01314-44
; Sequence 44, Application PC/TUS9601314
;  GENERAL INFORMATION:
;    APPLICANT:  M. Amin Arnaout
;    TITLE OF INVENTION:  METHODS FOR IDENTIFYING INTEGRIN
;    TITLE OF INVENTION:  ANTAGONISTS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/01314
;      FILING DATE:  30-JAN-96
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/380,167
;      FILING DATE:  30-JAN-95
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  John W. Freeman
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  00786/267001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1163
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
PCT-US96-01314-44

  Query Match             58.8%;  Score 3459;  DB 5;  Length 1163;
  Best Local Similarity   60.9%;  Pred. No. 1.2e-289;
  Matches  687;  Conservative  142;  Mismatches  294;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180

Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556
```

```
Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

Search completed: November 9, 2004, 12:21:04
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:37 ; Search time 27.25 Seconds
(without alignments)
4014.622 Million cell updates/sec

Title: US-09-902-481B-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5862	99.8	1153	1	RWHU1B	cell surface glyco
2	4470	76.1	1153	2	S00551	leukocyte surface
3	3479	59.2	1163	1	RWHU1C	cell surface glyco
4	1553.5	26.4	1170	2	S03308	cell surface glyco
5	1538.5	26.2	1163	2	I56126	lymphocyte fuction
6	1153	19.6	1179	2	A53213	integrin alpha-E c
7	1103.5	18.8	1151	2	A45226	integrin alpha-1 c
8	1085	18.5	1170	2	I45914	integrin alpha 2 s
9	1071	18.2	1178	2	S44142	VLA-2 protein homo
10	1068	18.2	1181	2	A33998	integrin alpha-2 c
11	1062	18.1	1180	2	A35854	integrin alpha-1 c
12	663	11.3	1039	2	A41131	lymphocyte-Peyer's
13	635	10.8	1038	2	S06046	integrin alpha-4 c
14	630	10.7	1035	2	I58409	integrin alpha-9 c
15	614.5	10.5	1041	2	T31437	integrin alpha cha
16	579.5	9.9	1054	2	JC7294	alphaP integrin -
17	572.5	9.7	1051	2	A35761	cell surface glyco
18	567.5	9.7	1053	2	I55534	VLA-3 alpha subuni
19	555.5	9.5	1053	2	S44250	integrin alpha-5 c
20	542.5	9.2	1034	2	A36108	integrin alpha-V c
21	535	9.1	1044	2	T10050	integrin alpha-v c
22	532	9.1	1049	2	A27079	fibronectin recept
23	532	9.1	1073	2	B36429	integrin alpha-6 c
24	531.5	9.0	1072	2	A38457	integrin alpha-6 c
25	530.5	9.0	1051	2	A40021	integrin alpha-3 c
26	526	9.0	1048	2	A27421	integrin alpha-5 c
27	525.5	8.9	1091	2	A41543	integrin alpha-6 c
28	517	8.8	1044	2	S16516	integrin alpha-8 c
29	496.5	8.4	1146	2	S40311	integrin - fruit f
30	496	8.4	1394	2	A29637	position-specific
31	495.5	8.4	1039	2	A34269	integrin alpha-2b
32	494	8.4	1137	2	JC5950	integrin alpha-7 c
33	490.5	8.3	1037	2	A60163	glycoprotein IIb -
34	488	8.3	1135	2	I61186	alpha-7 integrin -
35	486	8.3	126	2	B30892	leukocyte adhesion
36	470	8.0	1106	2	S38783	integrin alpha cha
37	467	7.9	1226	2	S44824	F54F2.1 protein -
38	453	7.7	1045	2	S60571	integrin alpha v c
39	445.5	7.6	1139	2	S28277	hypothetical prote
40	423	7.2	1115	2	T09403	integrin alpha cha
41	418.5	7.1	1115	2	T09433	integrin alpha cha
42	391	6.7	764	2	I36916	glycoprotein IIb -
43	309.5	5.3	1086	2	T18523	integrin alpha cha
44	299	5.1	604	2	I36917	glycoprotein IIb -
45	298.5	5.1	272	2	A55348	integrin alpha-1 -

ALIGNMENTS

RESULT 1
RWHU1B
cell surface glycoprotein CD11b precursor [validated] - human
N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Ma
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567
R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.
J. Biol. Chem. 263, 12403-12411, 1988
A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD
B.
A;Reference number: A31108; MUID:88315033; PMID:2457584
A;Accession: A31108
A;Molecule type: mRNA
A;Residues: 1-1153 <COR>
A;Cross-references: UNIPROT:P11215; GB:J03925; NID:g187284; PIDN:AAA59544.1; PID:g307148
A;Note: part of this sequence was confirmed by protein sequencing
R;Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
J. Cell Biol. 106, 2153-2158, 1988
A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor M
A;Reference number: A28915; MUID:88257215; PMID:2454931
A;Accession: A28915
A;Molecule type: mRNA
A;Residues: 1-499,501-965,'P',967-1153 <ARN>
A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AAA594
A;Note: the authors translated the codon TAC for residue 1129 as Thr
A;Note: part of this sequence, including the amino end of the mature protein, was confir
R;Shelley, C.S.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991
A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg
A;Reference number: A41600; MUID:92073318; PMID:1683702
A;Accession: A41600
A;Molecule type: DNA
A;Residues: 1-9 <SHE>
A;Cross-references: GB:M76724; NID:g180018; PIDN:AAA58410.1; PID:g553215
R;Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesi
A;Reference number: A94193; MUID:88190151; PMID:2833753
A;Accession: A30892
A;Molecule type: mRNA
A;Residues: 917-1042 <AR2>
A;Cross-references: GB:M18044
R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
A;Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor
A;Reference number: A32218; MUID:89098893; PMID:2563162
A;Accession: A32218
A;Molecule type: mRNA
A;Residues: 9-1153 <HIC>
A;Cross-references: GB:J04145; NID:g189068; PIDN:AAA59903.1; PID:g386975

```
A;Note: part of this sequence was confirmed by protein sequencing
R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A;Reference number: A46526; MUID:93123748; PMID:8419480
A;Accession: A46526
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-499,501-1153 <FLE>
A;Cross-references: GB:S52227; NID:g263047; PIDN:AAB24821.1; PID:g263049
A;Note: the last three bases of intron 13, CAG, are included in some but not all mature
A;Note: sequence extracted from NCBI backbone (NCBIP:121963)
R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A;Title: N-terminal sequence of human leukocyte glycoprotein Mo1: conservation across sp
A;Reference number: A90664; MUID:87076671; PMID:3539202
A;Accession: A26091
A;Molecule type: protein
A;Residues: 17-31 <PIE>
A;Experimental source: granulocytes
R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A;Title: Characterization of the myeloid-specific CD11b promoter.
A;Reference number: I52567; MUID:92144986; PMID:1346576
A;Accession: I52567
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-9 <RES>
A;Cross-references: GB:M84477; NID:g180184; PIDN:AAA51960.1; PID:g553219
C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C;Genetics:
A;Gene: GDB:ITGAM; CR3A
A;Cross-references: GDB:120599; OMIM:120980
A;Map position: 16p11.2-16p11.2
A;Note: promoter contains a GATA motif and two Sp1 consensus binding sites
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F;17-1108/Domain: extracellular #status predicted <EXT>
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;465-473/Region: calcium/magnesium binding #status predicted
F;530-538/Region: calcium/magnesium binding #status predicted
F;593-601/Region: calcium/magnesium binding #status predicted
F;1109-1134/Domain: transmembrane #status predicted <TMM>
F;1135-1153/Domain: intracellular #status predicted <INT>
F;86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

  Query Match             99.8%;  Score 5862;  DB 1;  Length 1153;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 2
S00551
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
N;Alternate names: complement-3 receptor alpha chain
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S00551; I59078
R;Pytela, R.
EMBO J. 7, 1371-1378, 1988
A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the
A;Reference number: S00551; MUID:88312584; PMID:3044779
A;Accession: S00551
A;Molecule type: DNA
A;Residues: 1-1153 <PYT>
A;Cross-references: UNIPROT:P05555; EMBL:X07640; NID:g52982; PIDN:CAA30479.1; PID:g5298
A;Note: the authors translated the codon CAC for residue 569 as Gln
```

R;Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts,
Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A;Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
A;Reference number: I59078; MUID:86287312; PMID:2942940
A;Accession: I59078
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 11-44 <RES>
A;Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
C;Genetics:
A;Gene: Mac-1
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C;Keywords: cell adhesion; glycoprotein; transmembrane protein
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1106-1129/Domain: transmembrane #status predicted <TMM>

Query Match 76.1%; Score 4470; DB 2; Length 1153;
Best Local Similarity 73.9%; Pred. No. 7.9e-303;
Matches 841; Conservative 146; Mismatches 149; Indels 2; Gaps 2;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||| |||||||:||||:|||| |: ||| |||| | || |:||||||| | ||
Db         17 FNLDTEHPMTFQENAKGFGQNVVQLGGTSVVVAAPQEAKAVNQTGALYQCDYSTSRCHPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               |||| |||||||||||| :| | |||||||||||| | ||||| |||:|||||| : ||:
Db         77 PLQVPPEAVNMSLGLSLAVSTVPQQLLACGPTVHQNCKENTYVNGLCYLFGSNLLRPPQQ 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              |||||| |||:|||  ||:|||||    ||:: |||:|||||| ||||||||||||:||
Db        137 FPEALRECPQQESDIVFLIDGSGSINNIDFQKMKEFVSTVMEQFKKSKTLFSLMQYSDEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              |||||| :|: ||:||| : || || ||| ||:|||||||||: |||||:|| |||::|
Db        197 RIHFTFNDFKRNPSPRSHVSPIKQLNGRTKTASGIRKVVRELFHKTNGARENAAKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||| |:|||||||| |||||||||:||   :||:||:|:|||| :||||::
Db        257 TDGEKFGDPLDYKDVIPEADRAGVIRYVIGVGNAFNKPQSRRELDTIASKPAGEHVFQVD 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||| ||||||:|||||||||||||:|||||||||||||:|||||||| :|||:||||
Db        317 NFEALNTIQNQLQEKIFAIEGTQTGSTSSFEHEMSQEGFSASITSNGPLLGSVGSFDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||||||:| |||| |||||||||||||||:|:|||||||||||||||||||| |||
Db        377 GAFLYTSKDKVTFINTTRVDSDMNDAYLGYASAVILRNRVQSLVLGAPRYQHIGLVVMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              :| | || : ::||:|||:|||||||||:|::|:|:|:|||||||||:|||||||||||
Db        437 ENFGTWEPHTSIKGSQIGSYFGASLCSVDMDADGNTNLILIGAPHYYEKTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| |||||||:|:|:|:|| ||||||||||||||||||||||||||||::|:|||||:|
Db        497 PRG-RARWQCEALLHGDQGHPWGRFGAALTVLGDVNGDKLTDVAIGAPGEQENQGAVYIF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              :| | : :| ||| || |:  || |||||||||||:|||||||:|| ||||||:||||:|
Db        556 YGASIASLSASHSHRIIGAHFSPGLQYFGQSLSGGKDLTMDGLMDLAVGAQGHLLLLRAQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||::| |||:|::|||:|| | :||:| |:|||||||| |:|:|:|||||| ||| ||
Db        616 PVLRLEATMEFSPKKVARSVFACQEQVLKNKDAGEVRVCLRVRKNTKDRLREGDIQSTVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||   || |:||||:|||:||| || | |||||| ||:|:::| ||||:||||:
Db        676 YDLALDPVRSRIRAFFDETKNNTRRRTQVFGLMQKCETLKLILPDCVDDSVSPIILRLNY 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || :||||||||| |||| |||:||||||||||:||||||||| | | || ||||
Db        736 TLVGEPLRSFGNLRPVLAMDAQRFFTAMFPFEKNCGNDSICQDDLSITMSAMGLDTLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRL-ACESASSTE 839
              ||::||::||:|||||||| |||| ::|  ||||| |  ||  ::: | :  ||:||:|
Db        796 GPQDFNMSVTLRNDGEDSYGTQVTVYYPSGLSYRKDSASQNPLTKKPWFVKPAESSSSSE 855

Qy        840 VSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTE 899
              ||||||: :|||||||| |||||||:|||||| || |||||||| | ||||| ||:||:
Db        856 GHGALKSTTWNINHPIFPANSEVTFNVTFDVDSHASFGNKLLLKAIVASENNMSRTHKTK 915

Qy        900 FQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              ||||||||||:||:|||   | :|||||||| ||:|:|||||  :|||||||||:|:|| :|
Db        916 FQLELPVKYAIYMIVTSDESSIRYLNFTASEMTSKVIQHQYQFNNLGQRSLPVSVVFWIP 975

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              |::|   :|| ||| ||:|||| |||::: | ||:|  :| : ||:|||:|||:|||||:
Db        976 VQINNVTVWDHPQVIFSQNLSSACHTEQKSPPHSNFRDQLERTPVLNCSVAVCKRIQCDL 1035

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | |  || || |||||||||||||||| |||:||: ||||||| | ||||| ::|||:||
Db       1036 PSFNTQEIFNVTLKGNLSFDWYIKTSHGHLLLVSSTEILFNDSAFALLPGQESYVRSKTE 1095

Qy       1080 TKVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||:|| ||:|||||||:|||:||||||| |||||||||||||||:|  |  | ||
Db       1096 TKVEPYEVHNPVPLIVGSSIGGLVLLALITAGLYKLGFFKRQYKDMMNEAAPQDAPPQ 1153
```

RESULT 3
RWHU1C
cell surface glycoprotein CD11c precursor - human
N;Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A36584; A35543; S00864
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A;Reference number: A36584
A;Contents: erratum
A;Accession: A36584
A;Molecule type: DNA
A;Residues: 1-1163 <COR>
A;Cross-references: UNIPROT:P20702
A;Note: this revision to the sequence from reference A35543 includes the carboxyl end
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.
A;Reference number: A35543; MUID:90153906; PMID:2303426
A;Accession: A35543
A;Molecule type: DNA
A;Residues: 1-834 <CO2>
A;Note: this sequence has been revised in reference A36584
R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A;Reference number: S00864; MUID:88166645; PMID:3327687
A;Accession: S00864
A;Molecule type: mRNA
A;Residues: 1-755,'L',757-1163 <CO3>
A;Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830
A;Note: part of this sequence was confirmed by protein sequencing
C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my
C;Genetics:
A;Gene: GDB:ITGAX; CD11C
A;Cross-references: GDB:119758; OMIM:151510
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C;Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F;20-1107/Domain: extracellular #status predicted <EXT>
F;149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F;1108-1133/Domain: transmembrane #status predicted <TMM>
F;1134-1163/Domain: intracellular #status predicted <INT>
F;61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match 59.2%; Score 3479; DB 1; Length 1163;
Best Local Similarity 61.1%; Pred. No. 9.3e-234;
Matches 690; Conservative 141; Mismatches 292; Indels 6; Gaps 4;

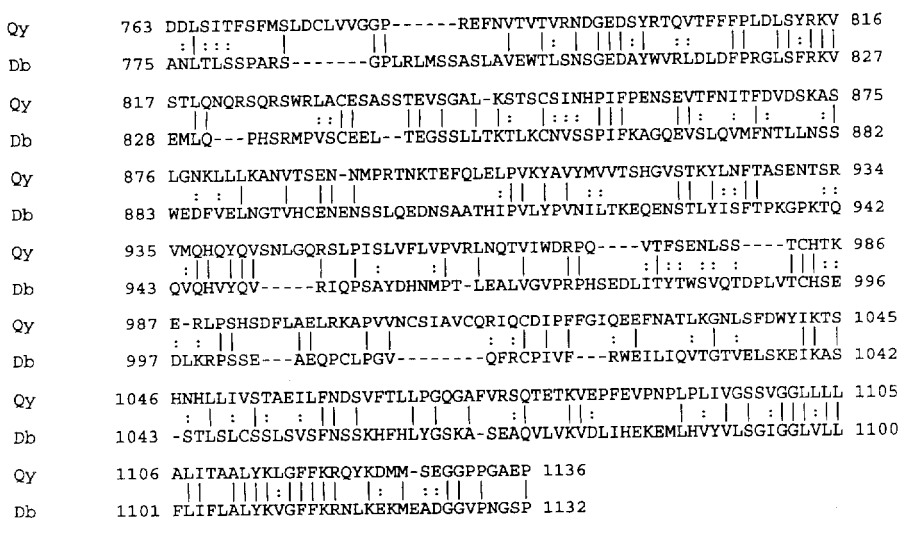
```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAADAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 4
S03308
cell surface glycoprotein CD11a precursor - human
N;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte functior
C;Species: Homo sapiens (man)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: S03308; A47458; A47565; A48759; S36044
R;Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.
J. Cell Biol. 108, 703-712, 1989
A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha subunit
A;Reference number: S03308; MUID:89139587; PMID:2537322
A;Accession: S03308
A;Molecule type: mRNA
A;Residues: 1-1170 <LAR>
A;Cross-references: UNIPROT:P20701; UNIPROT:Q9UBC8; EMBL:Y00796; NID:g31421; PIDN:CAA68
A;Note: part of this sequence was confirmed by protein sequencing
R;Cornwell, R.D.; Gollahon, K.A.; Hickstein, D.D.
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993
A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a) pro
A;Reference number: A47458; MUID:93248261; PMID:8097887
A;Accession: A47458
A;Molecule type: DNA
A;Residues: 1-20 <COR>
A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBIP:130863)
R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993
A;Title: Identification of cell-specific and developmentally regulated nuclear factors
A;Reference number: A47565; MUID:93281759; PMID:8099450
A;Accession: A47565
A;Molecule type: DNA
A;Residues: 1-20 <SHE>
A;Cross-references: GB:M95609
R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.
J. Biol. Chem. 268, 19305-19311, 1993
A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter.
A;Reference number: A48759; MUID:93374910; PMID:8103515
A;Accession: A48759
A;Molecule type: DNA
A;Residues: 1-20 <NUE>
A;Cross-references: EMBL:Z22804; NID:g311405; PIDN:CAA80461.1; PID:g311406
C;Genetics:
A;Gene: GDB:ITGAL; CD11A
A;Cross-references: GDB:119757; OMIM:153370
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen; tr
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted
F;154-317/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 26.4%; Score 1553.5; DB 2; Length 1170;
Best Local Similarity 34.7%; Pred. No. 1.5e-99;
Matches 407; Conservative 206; Mismatches 460; Indels 101; Gaps 37;

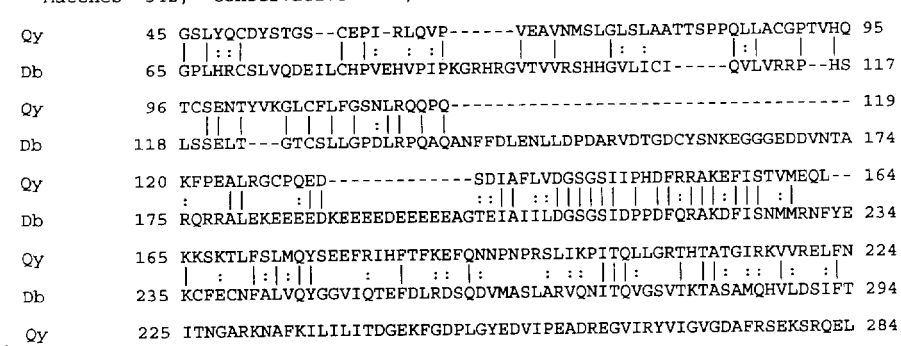
```
Qy          1 FNLDTENAMTFQ--ENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58
Db         26 YNLDVRGARSFSPPRAGRHFGYRVLQV-GNGVIVGAPGE---GNSTGSLYQCQSGTGHCL 81

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLR--- 115
Db         82 PVTLR-GSNYTSKYLGMTLATDPTDGSILACDPGLSRTCDQNTYLSGLCYLFRQNLQGPM 140

Qy        116 -QQPQKFPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLM 174
```

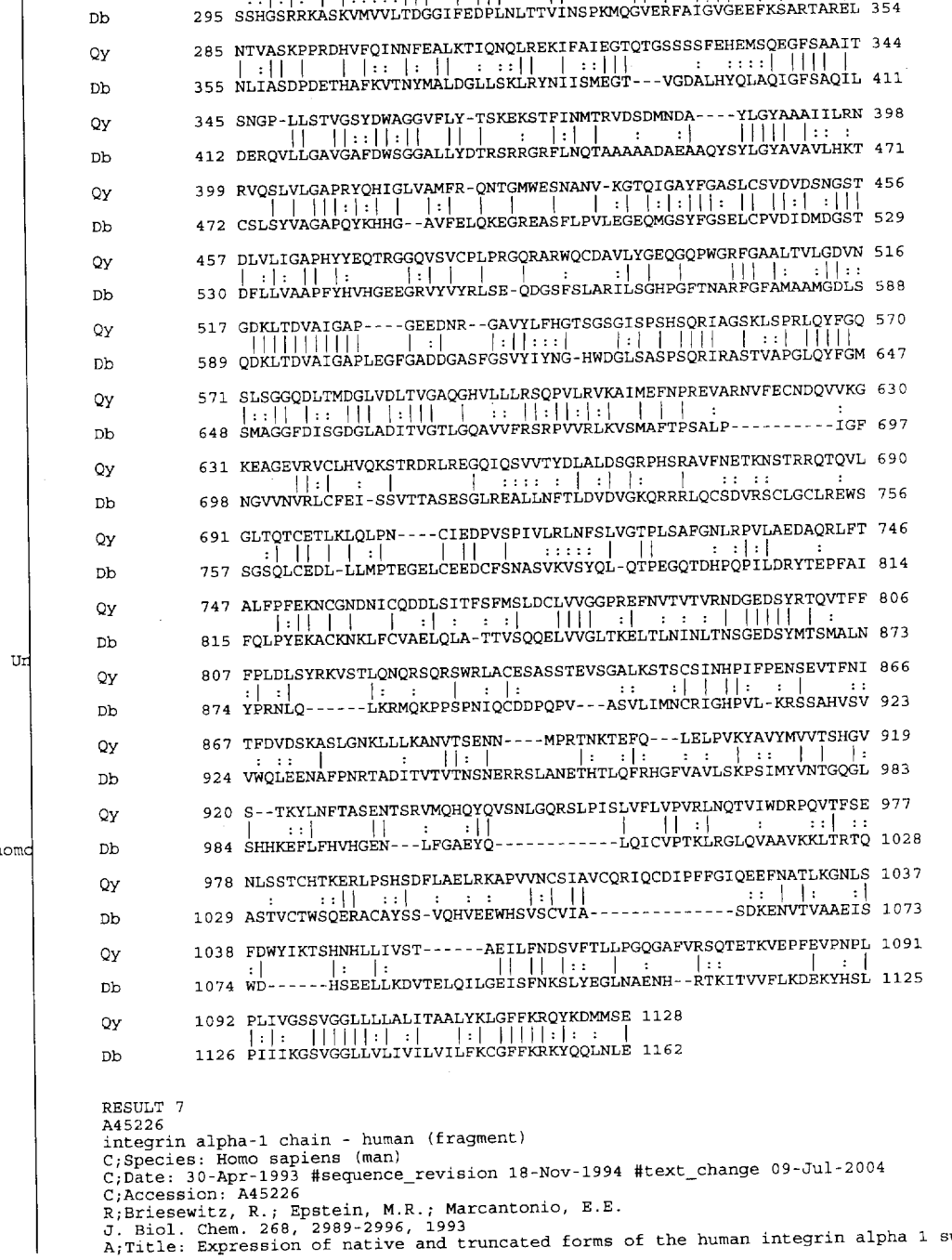
```
Db        141 LQGRPGFQECIKG----NVDLVFLFDGSMSLQPDEFQKILDFMKDVMKKLSNTSYQFAAV 196

Qy        175 QYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAF 234

Db        197 QFSTSYKTEFDFSDYVKWKDPDALLKHVKHMLLLTNTFGAINYVATEVFREELGARPDAT 256

Qy        235 KILILITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRD 294

Db        257 KVLIIITDGE--ATDSGNIDAAKD-----IIRYIIGIGKHFQTKESQETLHKFASKPASE 309

Qy        295 HVFQINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVG 354

Db        310 FVKILDTFEKLKDLFTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSRGHAVVGAVG 369

Qy        355 SYDWAGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQH 412

Db        370 AKDWAGGFLDLKADLQDDTFIGNEPLTPEVRAGYLGYTVTWLPSRQKTSLLASGAPRYQH 429

Qy        413 IGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQT 470

Db        430 MGRVLLFQEPQGGGHWSQVQTIHGTQIGSYFGGELCGVDVDQDGETELLLIGAPLFYGEQ 489

Qy        471 RGGQVSVCPLPRGQRARWQCDAV--LYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAP 528

Db        490 RGGRVFIY-----QRRQLGFEEVSELQGDPGYPLGRFGEAITALTDINGDGLVDVAVGAP 544

Qy        529 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTV 588

Db        545 LEE--QGAVYIFNGRHG-GLSPQPSQRIEGTQVLSGIQWFGRSIHGVKDLEGDGLADVAV 601

Qy        589 GAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVV-KGKEAGEVRVCLHVQKSTR 647

Db        602 GAESQMIVLSSRPVVDMVTLMSFSPAEIPVHEVECSYSTSNKMKEGVNITICFQI-KSLY 660

Qy        648 DRLREGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCI 707

Db        661 PQF-QGRLVANLTYTLQLDGHRTRRRGLFPGGRHELRRNIAVT-TSMSCTDFSFHFPVCV 718

Qy        708 EDPVSPIVLRLNFSL---VGTPLS--AFGN-----LRPVLAEDAQRLFTALFPFEKNCGN 757

Db        719 QDLISPINVSLNFSLWEEEGTPRDQRAQGKDIPPILRPSLHSETWEI-----PFEKNCGE 773

Qy        758 DNICQDDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVS 817

Db        774 DKKCEANLRVSFSPARSRALRLTAFASLSVELSLSNLEEDAYWVQLDLHFPPGLSFRKVE 833

Qy        818 TLQNQRSQRSWRLACES--ASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKAS 875

Db        834 ML---KPHSQIPVSCEELPEESRLLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNSS 887

Qy        876 LGNKLLLKANVTSENN----MPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASEN 931

Db        888 WGDSVELHANVTCNNEDSDLLEDNSATTI---IPILYPINILIQDQEDSTLYVSFTPKGP 944

Qy        932 TSRVMQHQYQV---SNLGQRSLP-ISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTK- 986

Db        945 KIHQVKHMYQVRIQPSIHDHNIPTLEAVVGVPQPPSEGPITHQWSVQMEPPV--PCHYED 1002

Qy        987 -ERLPSHSD--FLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIK 1043

Db       1003 LERLPDAAEPCLPGALFRCPVV----------------FRQEILVQVIGTLELVGEIE 1044

Qy       1044 TSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLL 1103

Db       1045 AS-SMFSLCSSLSISFNSSKHFHLYGSNASL-AQVVMKVDVVYEKQMLYLYVLSGIGGLL 1102

Qy       1104 LLALITAALYKLGFFKRQYKDMMSEG-GPPGAEP 1136

Db       1103 LLLLIFIVLYKVGFFKRNLKEKMEAGRGVPNGIP 1136

RESULT 5
I56126
lymphocyte fuction-associated molecule-1-alpha - mouse
C;Species: Mus musculus (house mouse)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: I56126
R;Kaufmann, Y.; Tseng, E.; Springer, T.A.
J. Immunol. 147, 369-374, 1991
A;Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit
A;Reference number: I56126; MUID:91268576; PMID:2051027
A;Accession: I56126
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1163 <RES>
A;Cross-references: UNIPROT:P24063; GB:M60778; NID:g198785; PIDN:AAA39426.1; PID:g19878
C;Genetics:
A;Gene: LFA-1
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
F;151-315/Domain: von Willebrand factor type A repeat homology <VWA1>

  Query Match             26.2%;  Score 1538.5;  DB 2;  Length 1163;
  Best Local Similarity   34.2%;  Pred. No. 1.6e-98;
  Matches  401;  Conservative  214;  Mismatches  458;  Indels   99;  Gaps   36;

Qy          1 FNLDTENAMTFQENA-RGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEP 59

Db         24 YNLDTRPTQSFLAQAGRHFGYQVLQIEDG-VVVGAPGE---GDNTGGLYHCRTSSEFCQP 79

Qy         60 IRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQ 119

Db         80 VSLH-GSNHTSKYLGMTLATDAAKGSLLACDPGLSRTCDQNTYLSGLCYLFPQSLEGPML 138

Qy        120 KFPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEE 179

Db        139 QNRPAYQECMKGKVDLVFLFDGSQSLDRKDFEKILEFMKDVMRKLSNTSYQFAAVQFSTD 198

Qy        180 FRIHFTFKEF-QNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILI 238

Db        199 CRTEFTFLDYVKQNKNPDVLLGSVQPMFLLTNTFRAINYVVAHVFKEESGARPDATKVLV 258

Qy        239 LITDGEKFGDPLGYEDVIPEADREG-------VIRYVIGVGDAFRSEKSRQELNTVASKP 291

Db        259 IITDG--------------EASDKGNISAAHDITRYIIGIGKHFVSVQKQKTLHIFASEP 304

Qy        292 PRDHVFQINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLS 351

Db        305 VEEFVKILDTFEKLKDLFTDLQRRIYAIEGTNRQDLTSFNMELSSSGISADLSKGHAVVG 364

Qy        352 TVGSYDWAGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPR 409

Db        365 AVGAKDWAGGFLDLREDLQGATFVGQEPLTSDVRGGYLGYTVAWMTSRSSRPLLAAGAPR 424

Qy        410 YQHIGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYY 467

Db        425 YQHVGQVLLFQAPEAGGRWNQTQKIEGTQIGSYFGGELCSVDLDQDGEAELLLIGAPLFF 484

Qy        468 EQTRGGQVSVCPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGA 527

Db        485 GEQRGGRVFTY---QRRQSLFEMVSELQGDPGYPLGRFGAAITALTDINGDRLTDVAVGA 541

Qy        528 PGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLT 587

Db        542 PLEE--QGAVYIFNGKPG-GLSPQPSQRIQGAQVFPGIRWFGRSIHGVKDLGGDRLADVV 598

Qy        588 VGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAG-EVRVCLHVQKST 646

Db        599 VGAEGRVVVLSSRPVVDVVTELSFSPEEIPVHEVECSYSAREEQKHGVKLKACFRIKPLT 658

Qy        647 RDRLREGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNC 706

Db        659 PQ--FQGRLLANLSYTLQLDGHRMRSRGLFPDGSHELSGNTSITP-DKSCLDFHFHFPIC 715

Qy        707 IEDPVSPIVLRLNFSLV---GTPLSAFGN-LRPVLAEDAQRLFTALFPFEKNCGNDNICQ 762

Db        716 IQDLISPINVSLNFSLLEEEGTPRDQKGRAMQPILRPSIHTV-TKEIPFEKNCGEDKKCE 774
```



RESULT 6
A53213
integrin alpha-E chain - human
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 09-Jul-2004
C;Accession: A53213
R;Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6016-6025, 1994
A;Title: Molecular cloning of the human mucosal lymphocyte integrin alpha(E) subunit. Un
A;Reference number: A53213; MUID:94164962; PMID:8119947
A;Accession: A53213
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1179 <SHA>
A;Cross-references: UNIPROT:P38570; GB:L25851; NID:g457244; PID:g457245
C;Genetics:
A;Gene: GDB:ITGAE
A;Cross-references: GDB:330801
A;Map position: 17p13
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
F;199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 19.6%; Score 1153; DB 2; Length 1179;
Best Local Similarity 29.1%; Pred. No. 1.2e-71;
Matches 342; Conservative 214; Mismatches 449; Indels 172; Gaps 39;





RESULT 7
A45226
integrin alpha-1 chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A45226
R;Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.
J. Biol. Chem. 268, 2989-2996, 1993
A;Title: Expression of native and truncated forms of the human integrin alpha 1 subuni

A;Reference number: A45226; MUID:93155124; PMID:8428973
A;Accession: A45226
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1151 <BRI>
A;Cross-references: UNIPROT:P56199
A;Experimental source: hepatoblastoma cell line HepG2
A;Note: sequence extracted from NCBI backbone (NCBIP:124326)
F;142-317/Domain: von Willebrand factor type A repeat homology <VWA1>

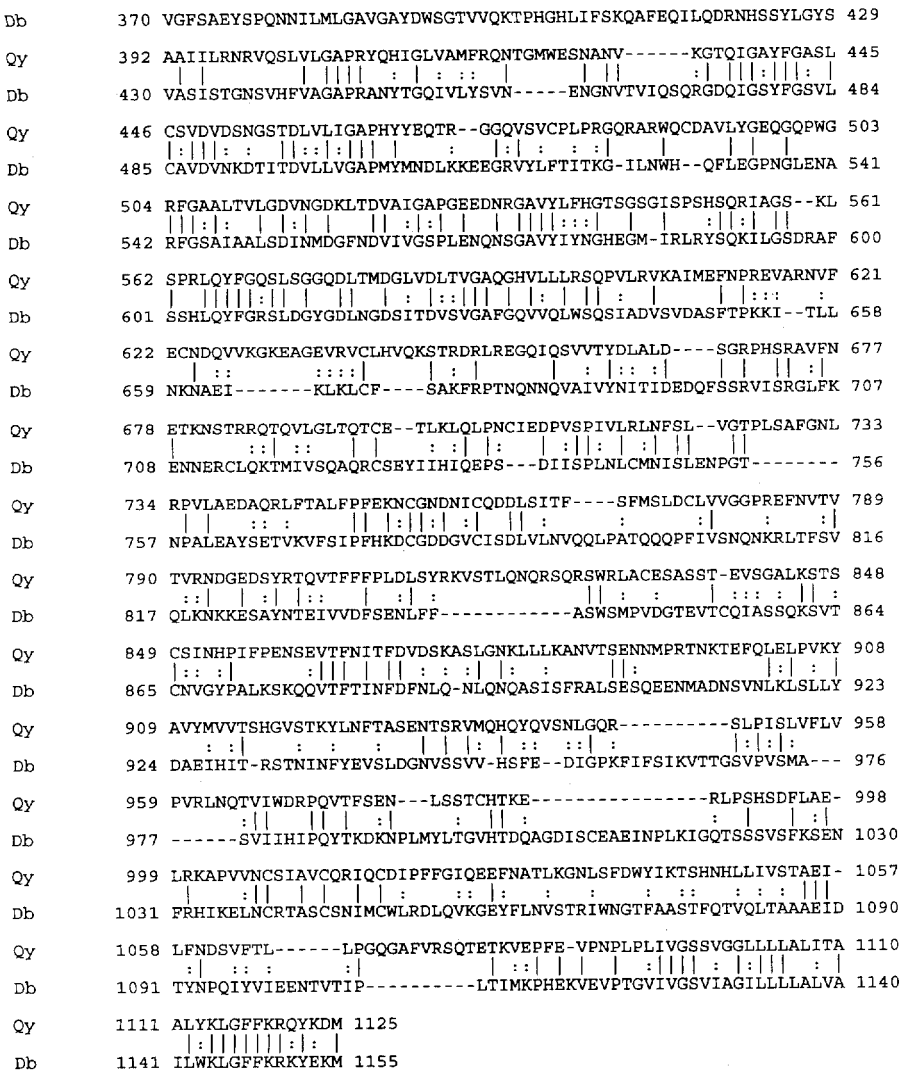
Query Match 18.8%; Score 1103.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 3.3e-68;
Matches 344; Conservative 212; Mismatches 486; Indels 195; Gaps 44;

```
Qy          1 FNLDTENAMTFQENARG-FGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQCDYSTGS  56
              ||:| :|:|||        || :| |    :|  |::|:|       |: | :|:|    |
Db          1 FNVDVKNSMTFSGPVEDMFGYTVQQYENEEGKWVLIGSPLVGQPKNRTGDVYKCPVGRGE  60

Qy         57 CEP-IRLQVPVEA---------VNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGL 106
               | ::| :||           ||: | :|  |       |||||      |    | |:
Db         61 SLPCVKLDLPVNTSIPNVTEVEKENMTFGSTL-VTNPNGGFLACGPLYAYRCGHLHYTTGI 119

Qy        107 CFLFGSNLRQQPQKFPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLK- 165
              |      :      |  :: |  :  ||  ::|||| || | |    |  |:: ::::
Db        120 CSDVSPTFQVVNSIAP--VQECSTQ-LDIVIVLDGSNSIYPWDSVTA--FLNDLLKRMDI 174

Qy        166 -KSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLGR-THTATGIRKVVRELF 223
               :|   ::|| |      |   :: :      | | |  || | || |     :| |
Db        175 GPKQTQVGIVQYGENVTHEFNLNKYSSTEEVLVAAKKIVQRGGRQTMTALGTDTARKEAF 234

Qy        224 NITNGARKNAFKILILITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFR-----SE 278
                 |||:   |:::::|||   |    : || : : | : |: | :  ::      :|
Db        235 TEARGARRGVKKVMVIVTDGESH-DNHRLKKVIQDCEDENIQRFSIAILGSYNRGNLSTE 293

Qy        279 KSRQELNTVASKPPRDHVFQINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEG 338
              |  :|: ::||:|   | | :::  || ||   | |:|||:| |   |::||| |||| |
Db        294 KFVEEIKSIASEPTEKHFFNVSDELALVTIVKTLGERIFALEATADQSAASFEMEMSQTG 353

Qy        339 FSAAITSNGPLLSTVGSYDWAGGVFLYTSKE-----KSTF-INMTRVDSDMNDAYLGYAA 392
              |||  : :  :|  ||:||| | | :  :      :|| :  |: :  :  :||||
Db        354 FSAHYSQDWVMLGAVGAYDWNGTVVMQKASQIIIPRNTTFNVESTKKNEPL-ASYLGYTV 412

Qy        393 AIILRNRVQSL-VLGAPRYQHIGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSVDVD 451
                    :  | : | ||| | | | ::|   |  :    : | |||:|||: | : |:|
Db        413 NSATASSGDVLYIAGQPRYNHTGQVIIYRMEDGNIKILQTLSGEQIGSYFGSILTTTDID 472

Qy        452 SNGSTDLVLIGAPHY---------------YEQTR-GGQVSVCPLPRGQRARWQCDAVLY 495
              : :||::|:||| |                 |||   |:|: |: :  :  | ::
Db        473 KDSNTDILLVGAPMYMGTEKEEQGKVYVYALNQTRFEYQMSLEPIKQTCCSSRQHNSCTT 532

Qy        496 GEQGQPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQ 554
              : :| | ||| |:  : |:| |   |: |||| |:|: ||||::|| ||  |   ::|
Db        533 ENKNEPCGARFGTAIAAVKDLNLDGFNDIVIGAPLEDDHGGAVYIYHG-SGKTIRKEYAQ 591

Qy        555 RIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIMEFNPR 614
              ||          |::||||: |  ||  ||| |:|:|   |   |  |: |  ||  | | |
Db        592 RIPSGGDGKTLKFFGQSIHGEMDLNGDGLTDVTIGGLGGAALFWSRDVAVVKVTMNFEPN 651

Qy        615 EVARNVFECNDQVVKGKEAG--EVRVCLHVQ-KSTRDRLREGQIQSVVTYDLALDSGRPH 671
              :|       |:   ::|||        ||  |: ||  | : |  :|    | : ||| |
Db        652 KVNIQKKNCH---MEGKETVCINATVCFEVKLKSKEDTIYEADLQ----YRVTLDSLRQI 704

Qy        672 SRAVFNET------KNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNFSLVGT 725
              ||: |: |       :| | |:::      |  :  :  :  :| |    : |:|:|
Db        705 SRSFFSGTQERKVQRNITVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDFNLT-D 755

Qy        726 PLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVGGPRE- 784
              | :      |||  :          || |:|||    |  |||:  :     | |:|    :
Db        756 PENG-----PVLDDSLPNSVHEYIPFAKDCGNKEKCISDLSLHVATTEKDLLIVRSQNDK 810

Qy        785 FNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEVSGAL 844
              |||:|:|:|  : :| |:     :  :| |  :  :  :|          :|||
Db        811 FNVSLTVKNTKDSAYNTRTIVHYSPNLVFSGIEAIQKD--------SCESN--------- 853

Qy        845 KSTSCSINHPIFPENSEVTFNITFDVDSKASLGN-KLLLKANVTSENNMPRTNKTEFQLE 903
              : :| : :|         |||  | |  ::   :  |  :| |   ||      :    :
Db        854 HNITCKVGYPFLRRGEMVTFKILFQFNTSYLMENVTIYLSATSDSEEPPETLSDNVVNIS 913

Qy        904 LPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSN------LGQRS-------L 950
              :|||| | :   |   |  :::  |:|    |:     : |      | ::|      |
Db        914 IPVKYEVGLQFYS-SASEYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSFPMPEL 972

Qy        951 PISLVF----------LVPVRLNQTVIWDRPQVTFSENLSSTCHTKE---------RLPS 991
              :|: |           | |  |:            |||  :   |  |          :: :
Db        973 KLSISFPNMTSNGYPVLYPTGLSS-----------SENANCRPHIFEDPFSINSGKKMTT 1021

Qy        992 HSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQE-----------------EFNATLK 1033
              :|    |::  :::|:       | |::   | :                    | |::
Db       1022 STD---HLKRGTILDCNTCKFATITCNLTSSDISQVNVSLILWKPTFIKSYFSSLNLTIR 1078

Qy       1034 GNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPL 1093
              | |          | |  |::|::                 :  :   ::   :|  :||
Db       1079 GEL-------RSENASLVLSSSN----------------QKRELAIQISKDGLPGRVPL 1114

Qy       1094 --IVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
                |: |:  |||||  |:  ||:|:|||||  |  | :
Db       1115 WVILLSAFAGLLLLMLLILALWKIGFFKRPLKKKMEK 1151
```

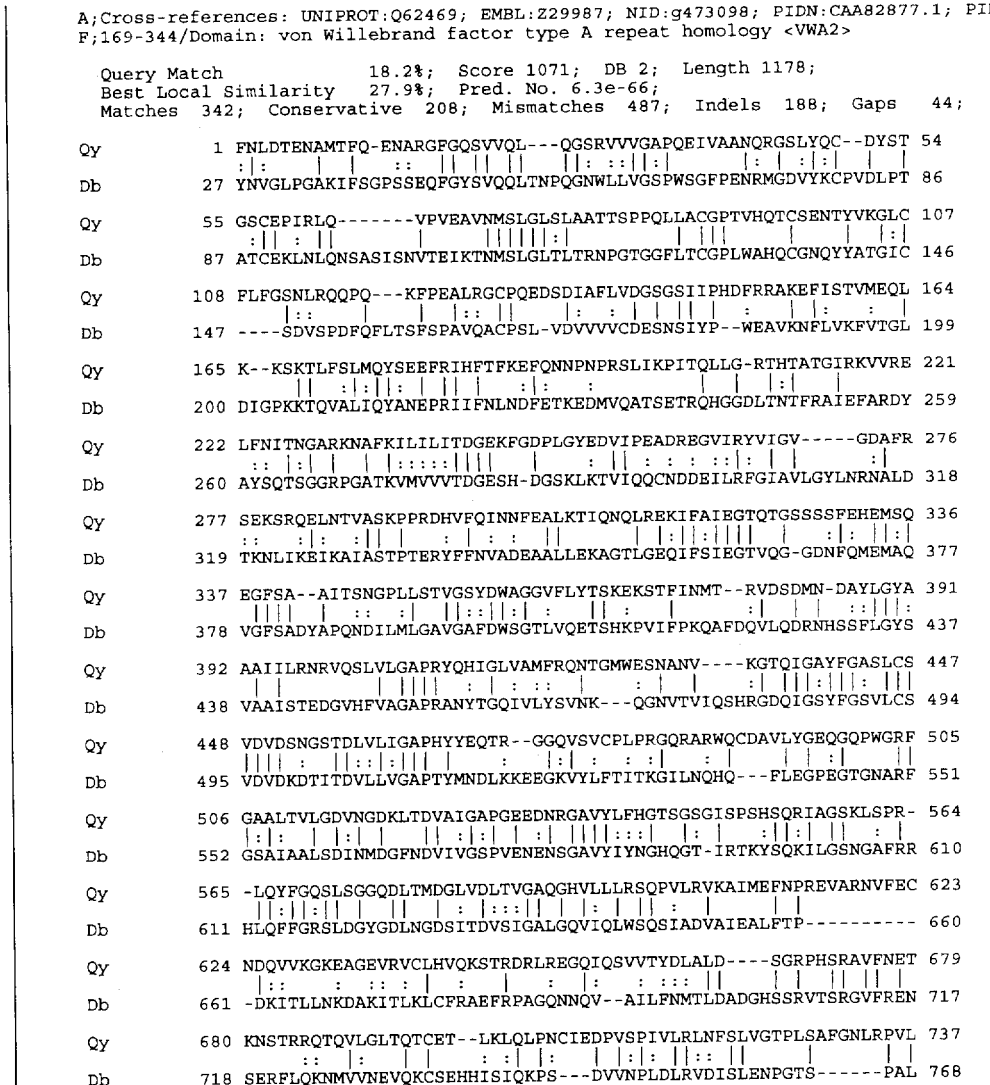
RESULT 8
I45914
integrin alpha 2 subunit - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C;Accession: I45914
R;Kamata, T.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A;Title: Identification of putative ligand binding sites within the I-domain of integrin
A;Reference number: A54402; MUID:94193647; PMID:7511592
A;Accession: I45914
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1170 <KAM>
A;Cross-references: UNIPROT:P53710; GB:L25886; NID:g439695; PIDN:AAB59255.1; PID:g439696
F;161-336/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.5%; Score 1085; DB 2; Length 1170;
Best Local Similarity 27.6%; Pred. No. 6.6e-67;
Matches 335; Conservative 217; Mismatches 495; Indels 168; Gaps 47;

```
Qy          1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQC--DYST  54
              :|:    |  |   ::  || :| |    :|: ::||:|      |: | :|:|  | ||
Db         19 YNVGLPKAKIFSGPSSEQFGYAVQQFINPKGNWLLVGSPWSGFPKNRMGDVYKCPVDLST  78

Qy         55 GSCEPIRLQ-------VPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLC 107
               :|| : ||       |      ||||||:|          | |||   | |     |:|
Db         79 TTCEKLNLQTSTSMSNVTEMKTNMSLGLTLTRNVGTGGFLTCGPLWAQQCGSQYYTTGVC 138

Qy        108 FLFGSNLRQQPQKFPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLK-- 165
              :  |    |  |:: ||    |:  : | | || | |    | |:   :: |
Db        139 SDVSPDF-QLRTSFAPAVQTCP-SFIDVVVVCDESNSIYPWD--AVKNFLEKFVQGLDIG 194

Qy        166 KSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLIKPITQLL----GRTHTATGIRKVVRE 221
              :||   |:||:   |: |     |::      :|| :|       |:|   |:
Db        195 PTKTQMGLIQYANNPRVVFNLNTFKSKD---EMIKATSQTFQYGGDLTNTFKAIQYARDT 251

Qy        222 LFNITNGARKNAFKILILITDGEKFGDPLGYEDVIPEADREGVIRYVIGV-----GDAFR 276
              ::   | |  | |:::::|||   |    : || : ::: ::|: | |       :|
Db        252 AYSTAAGGRPGATKVMVVVTDGESH-DGSKLKAVIDQCNKDNILRFGIAVLGYLNRNALD 310

Qy        277 SEKSRQELNTVASKPPRDHVFQINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQ 336
              ::  :|:  :|| |  | | :::   |      : |:||:|||| |   :|: ||||
Db        311 TKNLIKEIKAIASIPTERHFFNVSDEADLLEKAGTIGEQIFSIEGTVQG-GDNFQMEMSQ 369

Qy        337 EGFSAAIT--SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYA 391
              ||||  :  :|  :|  ||:|||:| |   |        |      ::  | |  :||||:
```



RESULT 9
S44142
VLA-2 protein homolog - mouse
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S44142
R;Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze
submitted to the EMBL Data Library, January 1994
A;Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not
A;Reference number: S44142
A;Accession: S44142
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1178 <EDE>

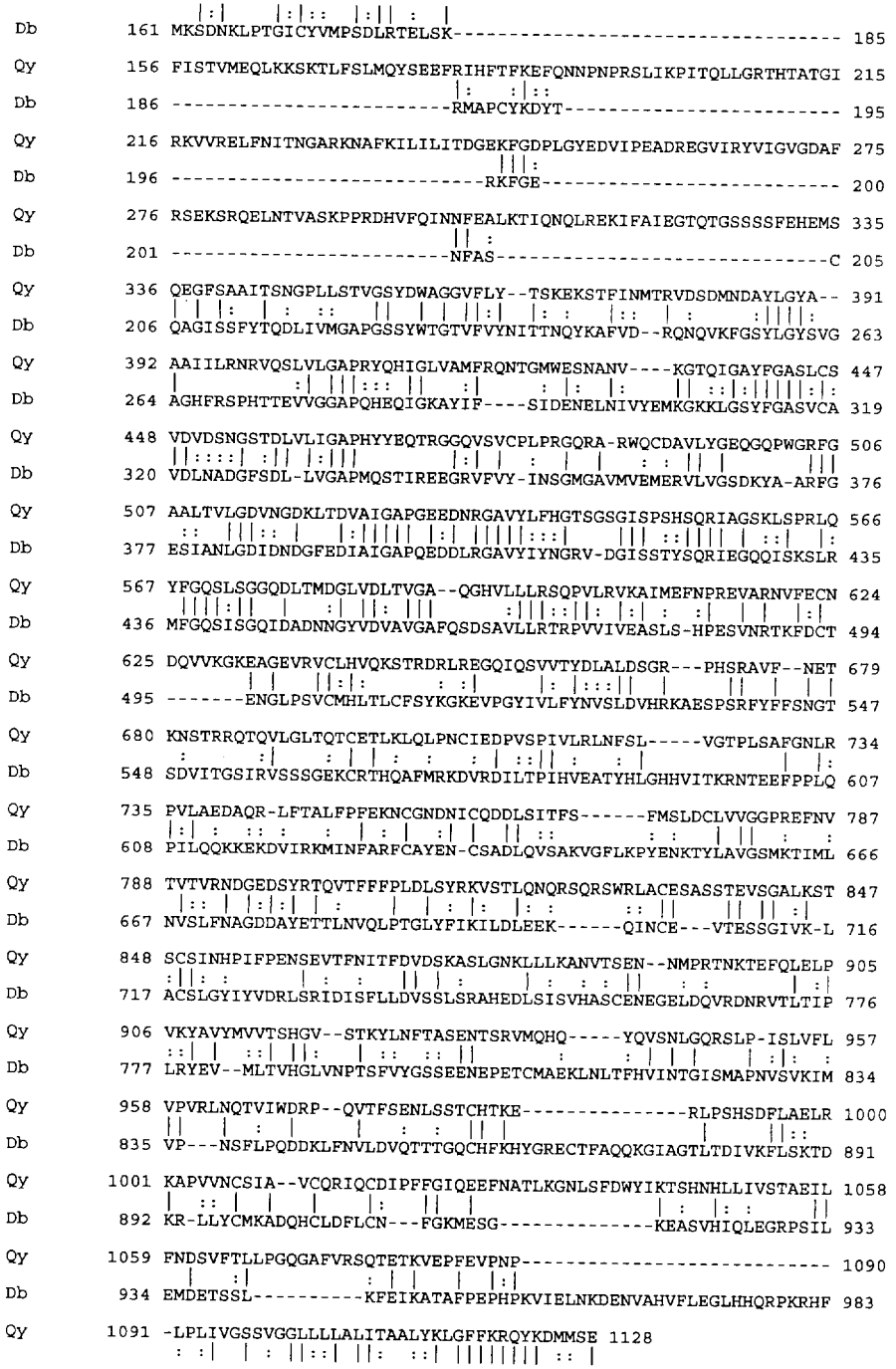


A;Cross-references: UNIPROT:Q62469; EMBL:Z29987; NID:g473098; PIDN:CAA82877.1; PID:g473
F;169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.2%; Score 1071; DB 2; Length 1178;
Best Local Similarity 27.9%; Pred. No. 6.3e-66;
Matches 342; Conservative 208; Mismatches 487; Indels 188; Gaps 44;

```
Db        929 LLYDAELHLT-RSTNINFYEISSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVSMA 984

Qy        956 FLV---------------------------------PVRLNQTVIWDRPQVTF-SENLS 980

Db        985 LVTIHIPQYTKEKNPLLYLTGIQTDQAGDISCTAEINPLKLPHTA----PSVSFKNENFR 1040

Qy        981 STCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDW 1040

Db       1041 ---HTKE----------------LDCRTTSCSNITCWLKDLHMKAEYFINVTTRVWNRT 1080

Qy       1041 YIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPPEVPNPLPLIVGSSVG 1100

Db       1081 FAASTFQTVQLTAAAEIDTHNPQLFVIEENAVTIPLMIMKPTEKAEVPT--GVIIGSIIA 1138

Qy       1101 GLLLLALITAALYKLGFFKRQYKDM 1125

Db       1139 GILLLLAMTAGLWKLGFFKRQYKKM 1163


RESULT 10
A33998
integrin alpha-2 chain precursor - human
N;Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 09-Jul-2004
C;Accession: A33998; B56793; A53117
R;Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet
A;Reference number: A33998; MUID:89308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: UNIPROT:P17301; GB:X17033; NID:g33906; PIDN:CAA34894.1; PID:g33907
A;Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R;Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIc, GPIc*, GPIIa an
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16,'V',18-21 <ZUT>
A;Cross-references: GB:L24121; NID:g400342; PIDN:AAA16619.2; PID:g4583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TMM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

  Query Match            18.2%;  Score 1068;  DB 2;  Length 1181;
  Best Local Similarity  26.9%;  Pred. No. 1e-65;
  Matches  329;  Conservative  214;  Mismatches  494;  Indels  188;  Gaps   43;

Qy          1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQC--DYST 54

Db         30 YNVGLPEAKIFSGPSSEQFGYAVQQFINPKGNWLLVGSPWSGFPENRMGDVYKCPVDLST 89

Qy         55 GSCEPIRLQ-------VPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLC 107

Db         90 ATCEKLNLQTSTSIPNVTEMKTNMSLGLILTRNMGTGGFLTCGPLWAQQCGNQYYTTGVC 149

Qy        108 FLFGSNLRQQPQKFPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLK-- 165

Db        150 SDISPDF-QLSASFSPATQPCPSL-IDVVVVCDESNSIYPWD--AVKNFLEKFVQGLDIG 205

Qy        166 KSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLG-RTHTATGIRKVVRELFN 224

Db        206 PTKTQVGLIQYANNPRVVFNLNTYKTKEEMIVATSQTSQYGGDLTNTFGAIQYARKYAYS 265

Qy        225 ITNGARKNAFKILILITDGEKFGDPLGYEDVIPEADREGVIRYVIGV-----GDAFRSEK 279

Db        266 AASGGRRSATKVMVVVTDGESH-DGSMLKAVIDQCNHDNILRFGIAVLGYLNRNALDTKN 324

Qy        280 SRQELNTVASKPPRDHVFQINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGF 339

Db        325 LIKEIKAIASIPTERYFFNVSDEAALLEKAGTLGEQIFSIEGTVQG-GDNFQMEMSQVGF 383

Qy        340 SAAITSNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYAAAI 394

Db        384 SADYSSQNDILMLGAVGAFGWSGTIVQKTSHGHLIFPKQAFDQILQDRNHSSYLGYSVAA 443

Qy        395 ILRNRVQSLVLGAPRYQHIGLVAMFRQNTGMWESNANV------KGTQIGAYFGASLCSV 448

Db        444 ISTGESTHFVAGAPRANYTGQIVLYSVN-----ENGNITVIQAHRGDQIGSYFGSVLCSV 498

Qy        449 DVDSNGSTDLVLIGAPHYYEQTR--GGQVSVCPLPRGQRARWQCDAVLYGEQGQPWGRFG 506

Db        499 DVDKDTITDVLLVGAPMYMSDLKKEEGRVYLFTIKKGILGQHQ---FLEGPEGIENTRFG 555

Qy        507 AALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGS--KLSPR 564

Db        556 SAIAALSDINMDGFNDVIVGSPLENQNSGAVYIYNGHQGT-IRTKYSQKILGSDGAFRSH 614

Qy        565 LQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECN 624

Db        615 LQYFGRSLDGYGDLNGDSITDVSIGAFGQVVQLWSQSIADVAIEASFTPEKI--TLVNKN 672

Qy        625 DQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVTYDLALD----SGRPHSRAVFNETK 680

Db        673 AQII-------LKLCF----SAKFRPTKQNNQVAIVYNITLDADGFSSRVTSRGLFKENN 721

Qy        681 NSTRRQTQVLGLTQTC--ETLKLQLPNCIEDPVSPIVLRLNFSLVGTPLSAFGNLRPVLA 738

Db        722 ERCLQKNMVVNQAQSCPEHIIYIQEPS---DVVNSLDLRVDISLENPGTS------PALE 772

Qy        739 EDAQRLFTALFPFEKNCGNDNICQDDLSITF----SFMSLDCLVVGGPREFNVTVTVRND 794

Db        773 AYSETAKVFSIPFHKDCGEDGLCISDLVLDVRQIPAAQEQPFIVSNQNKRLTFSVTLKNK 832

Qy        795 GEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASST-EVSGALKSTSCSINH 853

Db        833 RESAYNTGIVVDFSENLFF----------ASFSLPVDGTEVTCQVAASQKSVACDVGY 880

Qy        854 PIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEFQLELPVKYAVYMV 913

Db        881 PALKREQQVTFTINFDFNLQ-NLQNQASLSFQALSESQEENKADNLVNLKIPLLYDAEI- 938

Qy        914 VTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQR----------SLPISLVFLV----- 958

Db        939 ---HLTRSTNINFYEISSDGNVPSIVHSFEDVGPKFIFSLKVTTGSVPVSMATVIIHIPQ 995

Qy        959 ----------------------------PVRLNQTVIWDRPQVTF-SENLSSTCHTKER 988

Db        996 YTKEKNPLMYLTGVQTDKAGDISCNADINPLKIGQT----SSSVSFKSENFR---HTKE- 1047

Qy        989 LPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNH 1048

Db       1048 ---------------LNCRTASCSNVTCWLKDVHMKGEYFVNVTTRIWNGTFASSTFQT 1091

Qy       1049 LLIVSTAEI-LFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNP-----LP--LIVGSSVG 1100
```

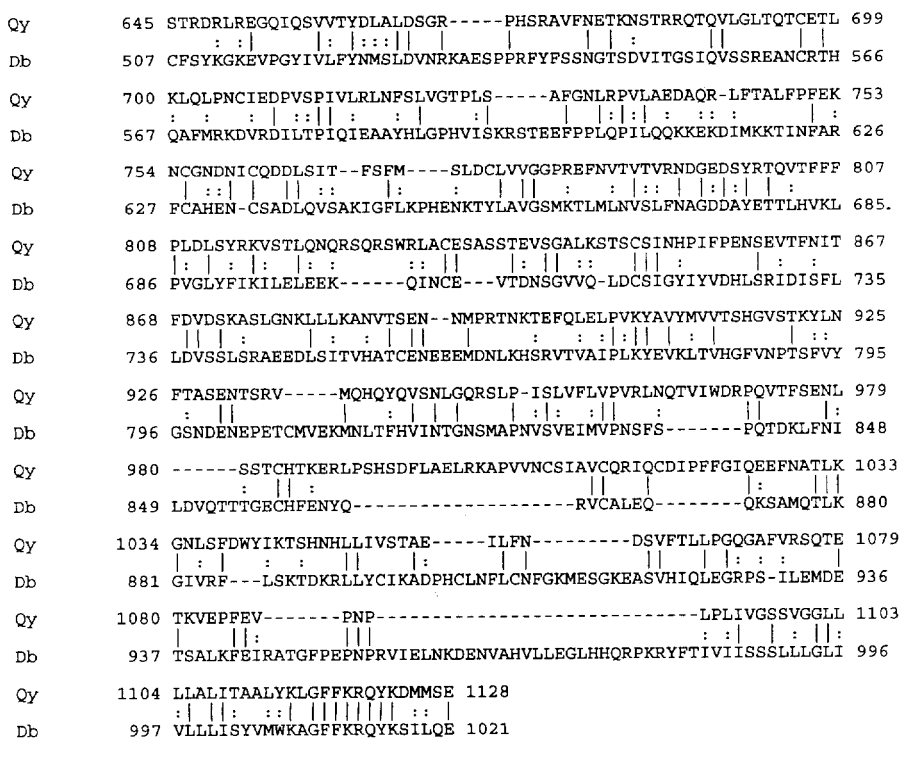
```
Db       1092 VQLTAAAEINTYNPEIYVI----------EDNTVTIPLMIMKPDEKAEVPTGVIIGSIIA 1141

Qy       1101 GLLLLALITAALYKLGFFKRQYKDM 1125
              |:|||  :  | |:|||||||:|: |
Db       1142 GILLLLALVAILWKLGFFKRKYEKM 1166


RESULT 11
A35854
integrin alpha-1 chain precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 09-Jul-2004
C;Accession: A35854; S11243
R;Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto,
J. Cell Biol. 111, 709-720, 1990
A;Title: Molecular cloning of the rat integrin alpha-1-subunit: a receptor for laminin a
A;Reference number: A35854; MUID:90338125; PMID:2380249
A;Accession: A35854
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1180 <IGN>
A;Cross-references: UNIPROT:P18614; GB:X52140; NID:g56493; PIDN:CAA36384.1; PID:g56494
C;Keywords: cell adhesion; cytoskeleton; transmembrane protein
F;170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

  Query Match             18.1%;  Score 1062;  DB 2;  Length 1180;
  Best Local Similarity   27.4%;  Pred. No. 2.7e-65;
  Matches  344;  Conservative  201;  Mismatches  478;  Indels  234;  Gaps   47;

Qy          1 FNLDTENAMTFQENARG-FGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQCDYSTGS 56
              ||:| :|:|:|        || :| |    :|  |::|:|       : | :|:|
Db         29 FNVDVKNSMSFSGPVEDMFGYTVQQYENEEGKWVLIGSPLVGQPKARTGDVYKCPVGRER 88

Qy         57 CEP-IRLQVPVEA---------VNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGL 106
                | ::| :||          ||: | :|  |      |||||      |    |  |:
Db         89 AMPCVKLDLPVNTSIPNVTEIKENMTFGSTL-VTNPNGGFLACGPLYAYRCGHLHYTTGI 147

Qy        107 CFLFGSNLRQQPQKFPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLK- 165
              |        :     |  :: |  :  ||  ::||| || |  :     |::  :::::
Db        148 CSDVSPTFQVVNSFAP--VQECSTQ-LDIVIVLDGSNSIYP--WESVIAFLNDLLKRMDI 202

Qy        166 -KSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLG-RTHTATGIRKVVRELF 223
               :|   ::|| |      |   :: :          | :  | :| || ||    :| |
Db        203 GPKQTQVGIVQYGENVTHEFNLNKYSSTEEVLVAANKIGRQGGLQTMTALGIDTARKEAF 262

Qy        224 NITNGARKNAFKILILITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFR-----SE 278
                |||:   |:::::|||    |     : || : : | : |: | :    :      :|
Db        263 TEARGARRGVKKVMVIVTDGESH-DNYRLKQVIQDCEDENIQRFSIAILGHYNRGNLSTE 321

Qy        279 KSRQELNTVASKPPRDHVFQINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEG 338
              |  :|: ::||:|   | | ::: || ||   | |:|||:| |   |::||| |||| |
Db        322 KFVEEIKSIASEPTEKHFFNVSDELALVTIVKALGERIFALEATADQSAASFEMEMSQTG 381

Qy        339 FSAAITSNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMND---AYLGYAAA 393
              |||  : :  :|  ||:||| | | :  : :     | |     : ||:   :||||
Db        382 FSAHYSQDWVMLGAVGAYDWNGTVVMQKANQMVIPHNTTFQTEPAKMNEPLASYLGYTVN 441

Qy        394 IILRNRVQSLVLGAPRYQHIGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSVDVDSN 453
                    : | ||| | | | ::: |       : | |||:|||: | ::|:| :
Db        442 SATIPGDVLYIAGQPRYNHTGQVVIYKMEDGNINILQTLGGEQIGSYFGSVLTTIDIDKD 501

Qy        454 GSTDLVLIGAPHY---------------YEQTR-GGQVSVCPLPRGQRARWQCDAVLYGE 497
                |||:|:||| |                |||   |:|: |: :   :  : ::
Db        502 SYTDLLLVGAPMYMGTEKEEQGKVYVYAVNQTRFEYQMSLEPIRQTCCSSLKDNSCTKEN 561

Qy        498 QGQPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQRI 556
              : :| | ||| |:  : |:| |   || |||| |:|: ||||::|| ||  |  :::|||
Db        562 KNEPCGARFGTAIAAVKDLNVDGFNDVVIGAPLEDDHAGAVYIYHG-SGKTIREAYAQRI 620

Qy        557 AGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIMEFNPREV 616
                    |::|||| |  ||  ||| |:|:|  |   |  :: |  ||  | | | :|
Db        621 PSGGDGKTLKFFGQSIHGEMDLNGDGLTDVTIGGLGGAALFWARDVAVVKVTMNFEPNKV 680

Qy        617 ARNVFECNDQVVKGKEAG--EVRVCLHVQ-KSTRDRLREGQIQSVVTYDLALDSGRPHSR 673
                    |   |:|||       :| ||: ||  | : |  :|    | : ||| |  ||
Db        681 NIQKKNCR---VEGKETVCINATMCFHVKLKSKEDSIYEADLQ----YRVTLDSLRQISR 733

Qy        674 AVFNET------KNSTRRQTQVLGLTQTCETLKLQLPNCI---------EDPVSPIVLRL 718
              : |: |       :| | |:::                ||          |    : : |
Db        734 SFFSGTQERKIQRNITVRESE-----------------CIRHSFYMLDKHDFQDSVRVTL 776

Qy        719 NFSLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLV 778
              :|:|   | :      ||| :          || |:|||   |  ||::  |      |:
Db        777 DFNLT-DPENG-----PVLDDALPNSVHEHIPFAKDCGNKERCISDLTLNVSTTEKSLLI 830

Qy        779 VGGPRE-FNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASS 837
              |    : |||::||:| |: : | |:       :| :  :  :|          :|||
Db        831 VKSQHDKFNVSLTVKNKGDSAYNTRTVVQHSPNLIFSGIEEIQKD--------SCESN-- 880

Qy        838 TEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLL-LKANVTSENNMPRTN 896
                  :: :| : :|       ||| | |  ::     | :: | |   || :   |
Db        881 -------QNITCRVGYPFLRAGETVTFKIIFQFNTSHLSENAIIHLSATSDSEEPLESLN 933

Qy        897 KTEFQLELPVKYAV----YMVVTSHGVST-------KYLNFTASENTSRVMQHQYQVSNL 945
                | : :|||| |     |   : | :|        :::| |  |:    :   | :
Db        934 DNEVNISIPVKYEVGLQFYSSASEHHISVAANETIPEFINST--EDIGNEINVFYTIRKR 991

Qy        946 GQRSLP---ISLVF----------LVPVRLNQTVIWD-------RP-------------Q 972
              |   :|   :| : |          | |:       |       ||             :
Db        992 GHFPMPELQLSISFPNLTADGYPVLYPIG------WSSSDNVNCRPRSLEDPFGINSGKK 1045

Qy        973 VTFS----------ENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI--- 1019
              :| |           :: ||||                          : ||:      | ::
Db       1046 MTISKSEVLKRGTIQDCSSTC----------------GVATITCSLLPSDLSQVNVSLL 1088

Qy       1020 ---PFFGIQEEF---NATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAF 1073
                  | | |:   |   | ||:| | |          | |  |  :|::
Db       1089 LWKPTF-IRAHFSSLNLTLRGELK-------SENSSLTLSSSN----------------- 1123

Qy       1074 VRSQTETKVEPFEVPNPLPL--IVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
                : :   ::    :|  :|| |: |:  |||||  |:  ||:|:|||||  |  | :
Db       1124 RKRELAIQISKDGLPGRVPLWVILLSAFAGLLLLMLLILALWKIGFFKRPLKKKMEK 1180


RESULT 12
A41131
lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
N;Alternate names: integrin alpha-4
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004
C;Accession: A41131; S16742
R;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.L.
J. Cell Biol. 115, 1149-1158, 1991
A;Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte
A;Reference number: A41131; MUID:92064645; PMID:1840602
A;Accession: A41131
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1039 <NEU>
A;Cross-references: UNIPROT:Q00651; EMBL:X53176; NID:g51484; PIDN:CAA37316.1; PID:g5148
C;Superfamily: integrin alpha-4 chain
C;Keywords: cytoskeleton; transmembrane protein

  Query Match             11.3%;  Score 663;  DB 2;  Length 1039;
  Best Local Similarity   22.7%;  Pred. No. 1.4e-37;
  Matches  281;  Conservative  187;  Mismatches  403;  Indels  368;  Gaps   48;

Qy          1 FNLDTENAMTFQ-ENARGFGQSVV-QLQGSR--VVVGAPQEIVAAN----QRGSLYQCDY 52
              :||| |||: :|  :    || |||     ||:  ::||||      :|      |::|:|
Db         41 YNLDPENALLYQGPSGTLFGYSVVLHSHGSKRWLIVGAPTASWLSNASVVNPGAIYRCGI 100

Qy         53 STG---SCEPIRLQVP----------VEAVNMSLGLSLAATTSP-PQLLACG---PTVHQ 95
                  :|| ::   |            |  | ||::|:        ::  ||     :
Db        101 RKNPNQTCEQLQSGSPSGEPCGKTCLEERDNQWLGVTLSRQPGENGSIVTCGHRWKNIFY 160

Qy         96 TCSENTYVKGLCFLFGSNLRQQPQKFPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKE 155
```



Db 984 TIIIITISLLLGLIVLLLISCVMWKAGFFKRQYKSILQE 1022

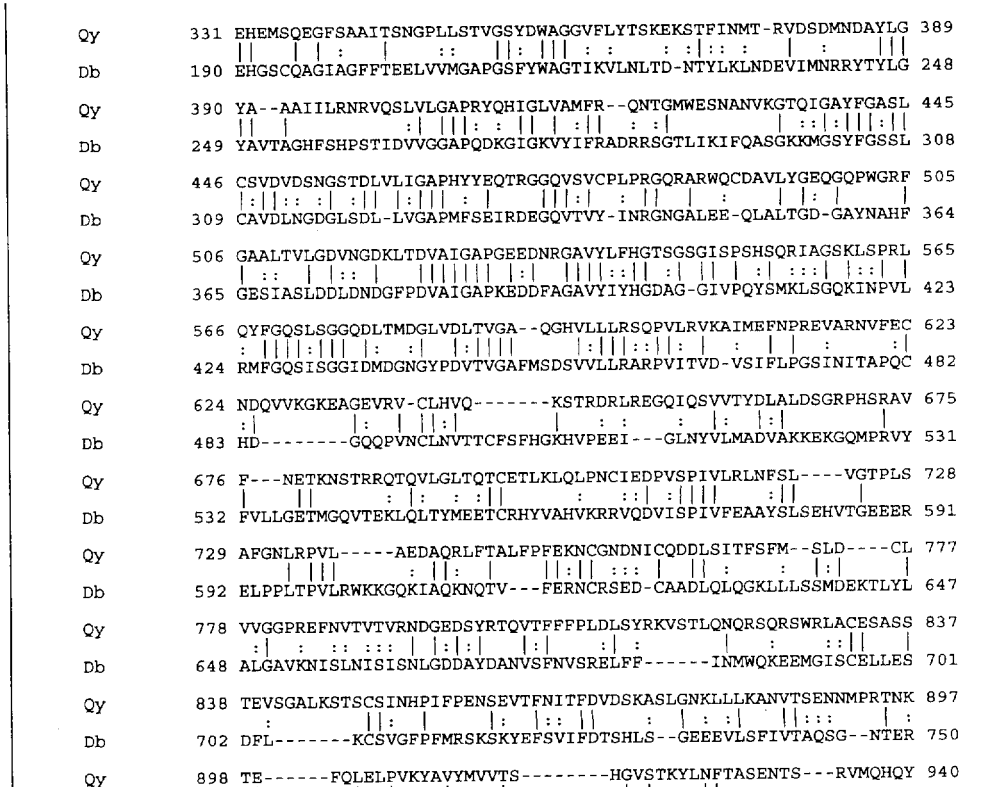
RESULT 13
S06046
integrin alpha-4 chain precursor - human
N;Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain
C;Species: Homo sapiens (man)
C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C;Accession: S06046; A39355; D28018
R;Takada, Y.; Elices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989
A;Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other inte
A;Reference number: S06046; MUID:89356603; PMID:2788572
A;Accession: S06046
A;Molecule type: mRNA
A;Residues: 1-1038 <TAK>
A;Cross-references: UNIPROT:P13612; GB:X16983; EMBL:X15356; NID:g33945; PIDN:CAA34852.1,
R;Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A;Title: Characterization of the alpha-4 integrin gene promoter.
A;Reference number: A39355; MUID:91239513; PMID:2034655
A;Accession: A39355
A;Molecule type: DNA
A;Residues: 1-93 <ROS>
A;Cross-references: GB:M62841
R;Takada, Y.; Strominger, J.L.; Hemler, M.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987
A;Title: The very late antigen family of heterodimers is part of a superfamily of molecu
A;Reference number: A94151; MUID:87204112; PMID:3033641
A;Accession: D28018
A;Molecule type: protein
A;Residues: 40-50,'E',52-53 <TA2>
C;Genetics:
A;Gene: GDB:ITGA4; CD49D
A;Cross-references: GDB:128032; OMIM:192975
A;Map position: 2q31-2q32
C;Superfamily: integrin alpha-4 chain
C;Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transme
F;1-39/Domain: signal sequence #status predicted <SIG>
F;40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.8%; Score 635; DB 2; Length 1038;
Best Local Similarity 25.3%; Pred. No. 1.2e-35;
Matches 249; Conservative 154; Mismatches 370; Indels 212; Gaps 39;



RESULT 14
I58409
integrin alpha-9 chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I58409; A49459
R;Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah
Oncogene 9, 611-619, 1994
A;Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small
A;Reference number: I58409; MUID:94119603; PMID:8290272
A;Accession: I58409
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1035 <RES>
A;Cross-references: UNIPROT:Q13797; GB:D25303; NID:g464180; PIDN:BAA04984.1; PID:g533327
R;Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.
J. Cell Biol. 123, 1289-1297, 1993
A;Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partne
A;Reference number: A49459; MUID:94064789; PMID:8245132
A;Accession: A49459
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 30-1035 <PAL>
A;Cross-references: GB:L24158
C;Superfamily: integrin alpha-4 chain
C;Keywords: glycoprotein; metal binding; transmembrane protein
F;1-27/Domain: signal sequence #status predicted <SIG>

Query Match 10.7%; Score 630; DB 2; Length 1035;
Best Local Similarity 26.5%; Pred. No. 2.7e-35;
Matches 237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;



RESULT 15
T31437
integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C;Species: Lytechinus variegatus (variegated urchin)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31437
R;Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A;Description: Alpha SU2, a sea urchin integrin which binds laminin.
A;Reference number: Z21035
A;Accession: T31437

```
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1041 <HER>
A;Cross-references: UNIPROT:O76378; EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC2
A;Experimental source: developmental stage embryo
C;Function:
A;Description: binds laminin
C;Superfamily: integrin alpha-2b chain

  Query Match           10.5%;  Score 614.5;  DB 2;  Length 1041;
  Best Local Similarity  24.8%;  Pred. No. 3.3e-34;
  Matches  236;  Conservative  155;  Mismatches  314;  Indels  245;  Gaps   41;

Qy        336 QEGFSAAITSNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDS----DMNDAYLG 389
Db        181 QAGFSGIIFSDNSALVMGAPGSYYLQGQIYVQSLLNRSV-VQATQESNTGTYSFDNSYRG 239

Qy        390 YAAAI--ILRNRVQSLVLGAPRYQHI-GLVAMFRQNTGMWESNANVKGTQIGAYFGASLC 446
Db        240 YSLALGDFNGDGVQDYVVGTPRAESLMGLVAIFDQNLNQFN---QVMGTQIVAYFGYSVT 296

Qy        447 SVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPLPRGQRARWQCDAVLYGEQ-------- 498
Db        297 VVDI-NNDTYDDLLVGAPMYMDDGPAIQ-----------RWEAGAVYVYLQNPDVGPGA 343

Qy        499 ------------GQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGS 546
Db        344 SNRLSLSSTLIGGQIRSRFGLSIASIGDSNQDGFNDVAIGAPYEGDDAGAVYIYHG-SAN 402

Qy        547 GISPSHSQRIAGSKLS-PRLQYFGQSLSGGQDLTMDGLVDLTVGAQ--GHVLLLRSQPVL 603
Db        403 GLKSTPAQVLTPSTLGHSGITTFGFSLQGGQDMDKNKYPDLLVGAESANTAVLIRTRPVV 462

Qy        604 RVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSV----- 658
Db        463 SLDATLNTEP-----------------------IGINLENKTYE-LADGTMVTSFIAMT 497

Qy        659 --------------VTYDLALDSG-RPHSRAVFNETKNSTRRQTQVLGL-TQTCETLKLQ 702
Db        498 CFTYTGNYLPDHIDISYTVTVDSGIIANRRAMFVDNDMSEITKTRRLAVSTQFCDPLRAY 557

Qy        703 LPNCIEDPVSPIVLRLNFSLVGTPLSAFGN---------LRPVLAEDAQRLFTALFPFEK 753
Db        558 VGNSIEDKLTPIKVTLQYDL--------NNDESRLQPHEILPIIDMATMSTQTKQVSIQN 609

Qy        754 NCGNDNICQDDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSY 813
Db        610 NCVN-NICIPDLDVTVT-PNLPNIVIGQTQELTLDVSLNNRGEDAFQSSLSVYYPLGLQF 667

Qy        814 RKVSTLQNQRSQRSWRLACESASSTEVSGALKSTSCSINHPIFPEN------SEVTFNIT 867
Db        668 VRL----ERKANMDFSVTCSEDSD------LRIITCDTGNPMVGKNILEFGLTLSTFQVS 717

Qy        868 FDVDS-----KASLGNKLLLKANVTSENNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTK 922
Db        718 GDKDSIEFYFKAESENS--EDPNTLENNELNMTVPVTVDCTLKLLSASYPEIVMYSTQED 775

Qy        923 YL-------NFTASENTSRVMQHQYQV-----SNLGQRSLPIS----------LVFLVPV 960
Db        776 YVVPPFPAKNASEADIGMEVM-HLYEVRNTGSSNAGEVSLNIQWPQKNEDGEYLFYLLGI 834

Qy        961 RLNQTVIWDRPQ---------------VTFSENLSSTCHTKERLPSHSDFLAELRKAPVV 1005
Db        835 MTEEGVTCQLTQGKANPEGVKLEPSTKAKLSNSTTQVSGRKRREPEVAEALAQTDN--VI 892

Qy       1006 NCSIAVCQRIQCDIPFFGIQEEFNAT------LKGNLSFDWYIKTSHNHLLIVSTAEILF 1059
Db        893 YCASDSCVLINCTI------DEINASKSKVVRILGRF---W--------------ERTF 928

Qy       1060 NDSVFTLLPGQGAFVRSQTETKVE--PFEVPNP-----------------LP------- 1092
Db        929 QKAVSELTPVVQATIASSASAAVKTIPYNIPLPRDFSDSTKASTLVTTEELVPPVTPIAW 988

Qy       1093 -LIVGSSVGGLLLLALITAALYKLGFFKR-------QYKDMMS--EGGPP 1132
Db        989 WIIVVSVLGGIILLLIIILGLWKCGFFERKKPGEEKEYAPVASADKDGPP 1038

Search completed: November  9, 2004, 12:19:34
Job time : 33.25 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:46:42 ; Search time 136.75 Seconds
(without alignments)
4783.919 Million cell updates/sec

Title: US-09-902-481B-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5846.5	99.5	1152	1	ITAM_HUMAN	P11215 homo sapien
2	4470	76.1	1153	1	ITAM_MOUSE	P05555 mus musculu
3	4381	74.6	1151	2	Q9JI30	Q9ji30 rattus norv
4	3809.5	64.8	920	2	Q28984	Q28984 sus scrofa
5	3479	59.2	1163	1	ITAX_HUMAN	P20702 homo sapien
6	3411	58.0	1162	1	ITAD_HUMAN	Q13349 homo sapien
7	3329.5	56.7	1188	2	Q6KAS4	Q6kas4 mus musculu
8	3329.5	56.7	1188	2	BAD21383	Bad21383 mus muscu
9	3320.5	56.5	1169	1	ITAX_MOUSE	Q9qxh4 mus musculu
10	3244.5	55.2	1161	1	ITAD_RAT	Q9qye7 rattus norv
11	1583	26.9	1165	1	ITAL_BOVIN	P61625 bos taurus
12	1557.5	26.5	1170	1	ITAL_HUMAN	P20701 homo sapien
13	1557	26.5	1166	2	Q6TYB8	Q6tyb8 bos taurus
14	1557	26.5	1166	2	AAQ90015	Aaq90015 bos tauru
15	1538.5	26.2	1163	1	ITAL_MOUSE	P24063 mus musculu
16	1534.5	26.1	1161	2	Q9WTV4	Q9wtv4 mus musculu
17	1524	25.9	1160	2	Q9R200	Q9r200 mus musculu
18	1409	24.0	1196	2	Q98TF1	Q98tf1 cyprinus ca
19	1359.5	23.1	1086	2	Q96HB1	Q96hb1 homo sapien
20	1358.5	23.1	1187	2	Q98TF0	Q98tf0 cyprinus ca
21	1278	21.7	927	2	Q8HZV0	Q8hzv0 bos taurus
22	1167.5	19.9	1167	2	O88340	O88340 rattus norv
23	1161.5	19.8	1167	1	ITAE_MOUSE	Q60677 mus musculu
24	1153	19.6	1179	1	ITAE_HUMAN	P38570 homo sapien
25	1123	19.1	1167	2	O88341	O88341 rattus norv
26	1103.5	18.8	1151	1	ITA1_HUMAN	P56199 homo sapien
27	1093.5	18.6	1189	1	ITAH_HUMAN	Q9ukx5 homo sapien
28	1085	18.5	1170	1	ITA2_BOVIN	P53710 bos taurus
29	1071	18.2	1178	1	ITA2_MOUSE	Q62469 mus musculu
30	1071	18.2	1178	2	Q6P1C7	Q6p1c7 mus musculu
31	1071	18.2	1178	2	AAH65139	Aah65139 mus muscu
32	1068	18.2	1181	1	ITA2_HUMAN	P17301 homo sapien
33	1064.5	18.1	1167	1	ITAG_HUMAN	O75578 homo sapien
34	1062	18.1	1180	1	ITA1_RAT	P18614 rattus norv
35	1053	17.9	1171	2	O42094	O42094 gallus gall
36	1052.5	17.9	1188	1	ITAH_MOUSE	P61622 mus musculu
37	1052.5	17.9	1188	2	Q7TQC3	Q7tqc3 mus musculu
38	1045	17.8	1038	2	Q8BS01	Q8bs01 mus musculu
39	1018.5	17.3	1160	2	Q8MKF4	Q8mkf4 felis silve
40	1018	17.3	895	2	Q9WUF8	Q9wuf8 mus sp. itg
41	865	14.7	348	2	Q8TES5	Q8tes5 homo sapien
42	850	14.5	1332	2	Q9BPQ8	Q9bpq8 halocynthia
43	808	13.8	205	2	Q63001	Q63001 rattus norv
44	780.5	13.3	304	2	Q6PG66	Q6pg66 mus musculu
45	780.5	13.3	304	2	AAH57200	Aah57200 mus muscu

ALIGNMENTS

RESULT 1
ITAM_HUMAN
ID ITAM_HUMAN STANDARD; PRT; 1152 AA.
AC P11215;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha
DE subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)
DE (Neutrophil adherence receptor).
GN Name=ITGAM; Synonyms=CR3A, CD11B;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88315033; PubMed=2457584;
RA Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;
RT "The human leukocyte adhesion glycoprotein Mac-1 (complement receptor
RT type 3, CD11b) alpha subunit. Cloning, primary structure, and relation
RT to the integrins, von Willebrand factor and factor B.";
RL J. Biol. Chem. 263:12403-12411(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88190151; PubMed=2833753;
RA Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;
RT "Molecular cloning of the alpha subunit of human and guinea pig
RT leukocyte adhesion glycoprotein Mo1: chromosomal localization and
RT homology to the alpha subunits of integrins.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88257215; PubMed=2454931;
RA Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;
RT "Amino acid sequence of the alpha subunit of human leukocyte adhesion
RT receptor Mo1 (complement receptor type 3).";
RL J. Cell Biol. 106:2153-2158(1988).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93123748; PubMed=8419480;
RA Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;
RT "Structural analysis of the CD11b gene and phylogenetic analysis of
RT the alpha-integrin gene family demonstrate remarkable conservation of
RT genomic organization and suggest early diversification during
RT evolution.";
RL J. Immunol. 150:480-490(1993).
RN [5]
RP SEQUENCE OF 9-1153 FROM N.A.
RX MEDLINE=89098893; PubMed=2563162;
RA Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L.,
RA Roth G.J.;
RT "cDNA sequence for the alpha M subunit of the human neutrophil
RT adherence receptor indicates homology to integrin alpha subunits.";

```
RL   Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).
RN   [6]
RP   SEQUENCE OF 1-9 FROM N.A.
RX   MEDLINE=92073318; PubMed=1683702;
RA   Shelley C.S., Arnaout M.A.;
RT   "The promoter of the CD11b gene directs myeloid-specific and
RT   developmentally regulated expression.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN   [7]
RP   SEQUENCE OF 1-9 FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=92144986; PubMed=1346576;
RA   Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT   "Characterization of the myeloid-specific CD11b promoter.";
RL   Blood 79:865-870(1992).
RN   [8]
RP   SEQUENCE OF 17-31.
RX   MEDLINE=87076671; PubMed=3539202;
RA   Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
RT   "N-terminal sequence of human leukocyte glycoprotein Mo1: conservation
RT   across species and homology to platelet IIb/IIIa.";
RL   Biochim. Biophys. Acta 874:368-371(1986).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX   MEDLINE=95171458; PubMed=7867070;
RA   Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT   "Crystal structure of the A domain from the alpha subunit of integrin
RT   CR3 (CD11b/CD18).";
RL   Cell 80:631-638(1995).
RN   [10]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX   MEDLINE=96363671; PubMed=8747460;
RA   Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
RT   "Two conformations of the integrin A-domain (I-domain): a pathway for
RT   activation?";
RL   Structure 3:1333-1340(1995).
RN   [11]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX   MEDLINE=98362595; PubMed=9687375;
RA   Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA   Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
RA   Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
RA   Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT   "Cation binding to the integrin CD11b I domain and activation model
RT   assessment.";
RL   Structure 6:923-935(1998).
RN   [12]
RP   3D-STRUCTURE MODELING OF 17-616.
RX   MEDLINE=98226734; PubMed=9560195;
RA   Oxvig C., Springer T.A.;
RT   "Experimental support for a beta-propeller domain in integrin alpha-
RT   subunits and a calcium binding site on its lower surface.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC   -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various
CC       adhesive interactions of monocytes, macrophages and granulocytes
CC       as well as in mediating the uptake of complement-coated particles.
CC       It is identical with CR-3, the receptor for the iC3b fragment of
CC       the third complement component. It probably recognizes the R-G-D
CC       peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for
CC       fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides
CC       of fibrinogen gamma chain.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-M
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC       granulocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
```


```
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

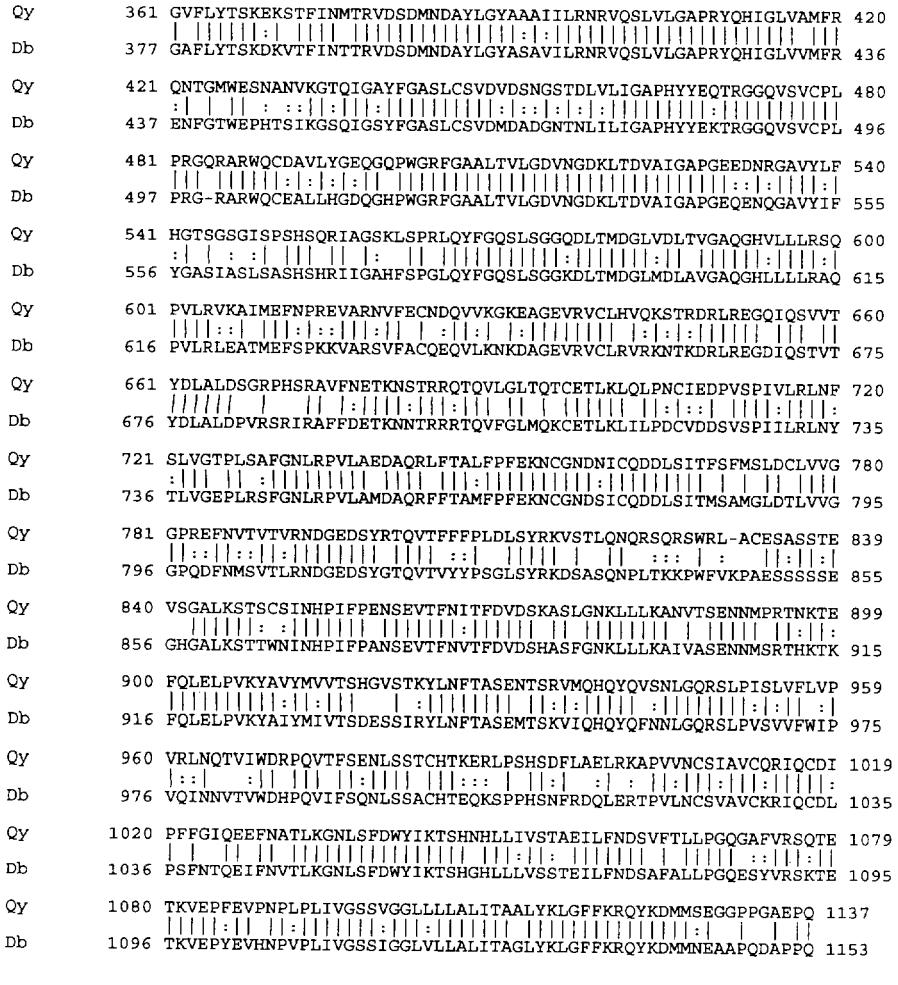

```
DR   EMBL; J03925; AAA59544.1; -.
DR   EMBL; M18044; AAA59491.1; -.
DR   EMBL; J04145; AAA59903.1; -.
DR   EMBL; S52227; AAB24821.1; -.
DR   EMBL; S52152; AAB24821.1; JOINED.
DR   EMBL; S52153; AAB24821.1; JOINED.
DR   EMBL; S52154; AAB24821.1; JOINED.
DR   EMBL; S52155; AAB24821.1; JOINED.
DR   EMBL; S52157; AAB24821.1; JOINED.
DR   EMBL; S52159; AAB24821.1; JOINED.
DR   EMBL; S52161; AAB24821.1; JOINED.
DR   EMBL; S52164; AAB24821.1; JOINED.
DR   EMBL; S52165; AAB24821.1; JOINED.
DR   EMBL; S52167; AAB24821.1; JOINED.
DR   EMBL; S52169; AAB24821.1; JOINED.
DR   EMBL; S52170; AAB24821.1; JOINED.
DR   EMBL; S52173; AAB24821.1; JOINED.
DR   EMBL; S52174; AAB24821.1; JOINED.
DR   EMBL; S52180; AAB24821.1; JOINED.
DR   EMBL; S52181; AAB24821.1; JOINED.
DR   EMBL; S52184; AAB24821.1; JOINED.
DR   EMBL; S52189; AAB24821.1; JOINED.
DR   EMBL; S52191; AAB24821.1; JOINED.
DR   EMBL; S52192; AAB24821.1; JOINED.
DR   EMBL; S52203; AAB24821.1; JOINED.
DR   EMBL; S52212; AAB24821.1; JOINED.
DR   EMBL; S52213; AAB24821.1; JOINED.
DR   EMBL; S52216; AAB24821.1; JOINED.
DR   EMBL; S52219; AAB24821.1; JOINED.
DR   EMBL; S52220; AAB24821.1; JOINED.
DR   EMBL; S52221; AAB24821.1; JOINED.
DR   EMBL; S52222; AAB24821.1; JOINED.
DR   EMBL; S52226; AAB24821.1; JOINED.
DR   EMBL; M76724; AAA58410.1; -.
DR   EMBL; M84477; AAA51960.1; -.
DR   PIR; A31108; RWHU1B.
DR   PDB; 1A8X; Model; @=17-1152.
DR   PDB; 1BHO; X-ray; 1/2=-.
DR   PDB; 1BHQ; X-ray; 1/2=-.
DR   PDB; 1IDN; X-ray; 1/2=-.
DR   PDB; 1IDO; X-ray; @=140-331.
DR   PDB; 1JLM; X-ray; @=143-334.
DR   PDB; 1M1U; X-ray; A=137-331.
DR   PDB; 1MF7; X-ray; A=144-337.
DR   PDB; 1N9Z; X-ray; A=140-335.
DR   PDB; 1NA5; X-ray; A=144-345.
DR   Genew; HGNC:6149; ITGAM.
DR   MIM; 120980; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0007155; P:cell adhesion; TAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
KW   Glycoprotein; Integrin; Magnesium; Receptor; Repeat; Signal;
KW   Transmembrane.
```

```
FT   SIGNAL        1     16
FT   CHAIN        17   1152       Integrin alpha-M.

  Query Match             99.5%;  Score 5846.5;  DB 1;  Length 1152;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||| |||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||::||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 2
ITAM_MOUSE
ID   ITAM_MOUSE     STANDARD;      PRT;  1153 AA.
AC   P05555; Q8CA73;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha
DE   subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).
GN   Name=Itgam;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=88312584; PubMed=3044779;
RA   Pytela R.;
RT   "Amino acid sequence of the murine Mac-1 alpha chain reveals homology
RT   with the integrin family and an additional domain related to von
RT   Willebrand factor.";
RL   EMBO J. 7:1371-1378(1988).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [3]
RP   SEQUENCE OF 11-45 FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Spleen;
RX   MEDLINE=86287312; PubMed=2942940;
RA   Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E.,
RA   Larson R.S., Roberts T.M., Springer T.A.;
RT   "A partial genomic DNA clone for the alpha subunit of the mouse
```

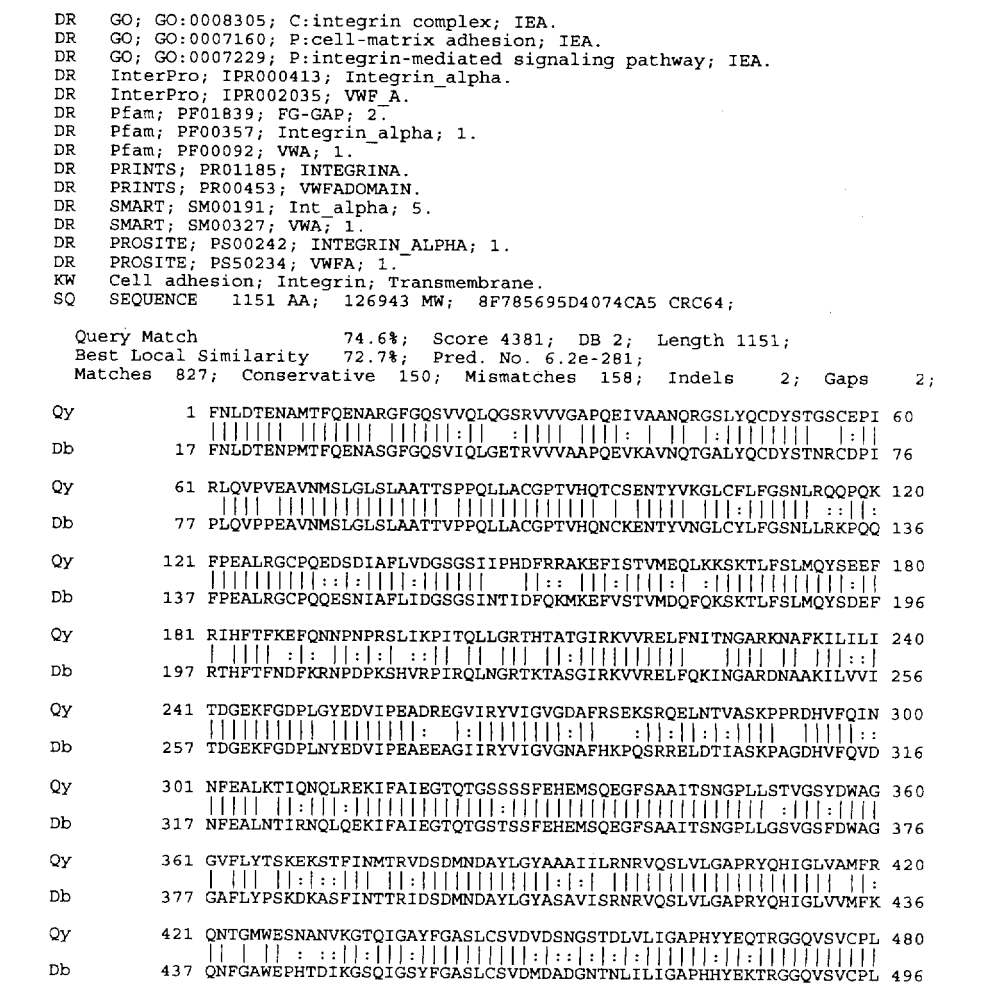
```
RT   complement receptor type 3 and cellular adhesion molecule Mac-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).
RN   [4]
RP   SEQUENCE OF 17-28.
RX   MEDLINE=85188276; PubMed=3887182;
RA   Springer T.A., Teplow D.B., Dreyer W.J.;
RT   "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
RT   glycoproteins and unexpected relation to leukocyte interferon.";
RL   Nature 314:540-542(1985).
CC   -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various
CC       adhesive interactions of monocytes, macrophages and granulocytes
CC       as well as in mediating the uptake of complement-coated particles.
CC       It is identical with CR-3, the receptor for the iC3b fragment of
CC       the third complement component. It probably recognizes the R-G-D
CC       peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for
CC       fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides
CC       of fibrinogen gamma chain. Alpha-M/beta-2 play a critical role in
CC       mast cell development and in immune complex-mediated
CC       glomerulonephritis. Mice expressing a null mutation of the alpha-M
CC       subunit gene demonstrate increase in neutrophil accumulation, in
CC       response to a impaired degranulation and phagocytosis, events that
CC       apparently accelerate apoptosis in neutrophils. These mice develop
CC       obesity.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta chain. Alpha-M
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P05555-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P05555-2; Sequence=VSP_010473;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC       granulocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X07640; CAA30479.1; -.
DR   EMBL; AK039444; BAC30350.1; -.
DR   EMBL; M14293; AAA39484.1; -.
DR   PIR; S00551; S00551.
DR   HSSP; P11215; 1BHQ.
DR   MGD; MGI:96607; Itgam.
DR   GO; GO:0009897; C:external side of plasma membrane; IDA.
DR   GO; GO:0007155; P:cell adhesion; IMP.
DR   GO; GO:0045123; P:cellular extravasation; IMP.
DR   GO; GO:0030593; P:neutrophil chemotaxis; IMP.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Alternative splicing; Calcium; Cell adhesion;
KW   Direct protein sequencing; Glycoprotein; Integrin; Receptor; Repeat;
KW   Signal; Transmembrane.
```

```
FT   SIGNAL       1    16
FT   CHAIN       17  1153       Integrin alpha-M.
FT   DOMAIN      17  1105       Extracellular (Potential).
FT   TRANSMEM  1106  1129       Potential.
FT   DOMAIN    1130  1153       Cytoplasmic (Potential).
FT   REPEAT      31    84       FG-GAP 1.
FT   REPEAT      85   163       FG-GAP 2.
FT   DOMAIN     164   350       VWFA.
FT   REPEAT     337   400       FG-GAP 3.
FT   REPEAT     401   452       FG-GAP 4.
FT   REPEAT     454   515       FG-GAP 5.
FT   REPEAT     517   575       FG-GAP 6.
FT   REPEAT     580   632       FG-GAP 7.
FT   CA_BIND    465   473       Potential.
FT   CA_BIND    529   537       Potential.
FT   CA_BIND    592   600       Potential.
FT   SITE      1132  1136       GFFKR motif.
FT   DISULFID    66    73       By similarity.
FT   DISULFID   105   123       By similarity.
FT   DISULFID   654   711       By similarity.
FT   DISULFID   770   776       By similarity.
FT   DISULFID   999  1023       By similarity.
FT   DISULFID  1028  1033       By similarity.
FT   CARBOHYD    58    58       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    86    86       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   391   391       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   696   696       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   734   734       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   772   772       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   801   801       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   881   881       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   907   907       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   941   941       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   980   980       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   994   994       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1022  1022       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1045  1045       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1051  1051       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1076  1076       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC   453   569       Missing (in isoform 2).
FT                              /FTId=VSP_010473.
FT   CONFLICT    37    37       N -> S (in Ref. 2).
FT   CONFLICT   683   683       V -> G (in Ref. 2).
SQ   SEQUENCE  1153 AA;  127480 MW;  178DB988AECB0343 CRC64;
```

```
  Query Match             76.1%;  Score 4470;  DB 1;  Length 1153;
  Best Local Similarity   73.9%;  Pred. No. 7.9e-287;
  Matches  841;  Conservative  146;  Mismatches  149;  Indels    2;  Gaps    2;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||: ||||||||:||||:|||| |: ||| |||| | || |:||||||||  | ||
Db         17 FNLDTEHPMTFQENAKGFGQNVVQLGGTSVVVAAPQEAKAVNQTGALYQCDYSTSRCHPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               |||| |||||||||||| :| | |||||||||| | ||||| |||:|||||| : ||:
Db         77 PLQVPPEAVNMSLGLSLAVSTVPQQLLACGPTVHQNCKENTYVNGLCYLFGSNLLRPPQQ 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              |||||| |||::||| ||:||||||    ||:: |||:|||||| ||||||||||||:||
Db        137 FPEALRECPQQESDIVFLIDGSGSINNIDFQKMKEFVSTVMEQFKKSKTLFSLMQYSDEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              |||||| :|: ||:||| : || || ||| ||:|||||||||: |||||:|| |||::|
Db        197 RIHFTFNDFKRNPSPRSHVSPIKQLNGRTKTASGIRKVVRELFHKTNGARENAAKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||| |:|||||||| ||||||||||:||    :||:||:|:|||| :||||::
Db        257 TDGEKFGDPLDYKDVIPEADRAGVIRYVIGVGNAFNKPQSRRELDTIASKPAGEHVFQVD 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||| |||||:|||||||||||||:|||||||||||||:|||||||| :|||:||||
Db        317 NFEALNTIQNQLQEKIFAIEGTQTGSTSSFEHEMSQEGFSASITSNGPLLGSVGSFDWAG 376
```



```
RESULT 3
Q9JI30
ID   Q9JI30      PRELIMINARY;      PRT;  1151 AA.
AC   Q9JI30;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Integrin beta 2 alpha subunit.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fathallah D.M. Sr., Zerria K. Jr.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR   EMBL; AF268593; AAF81280.1; -.
DR   HSSP; P11215; 1BHQ.
```



```
DR   GO; GO:0008305; C:integrin complex; IEA.
DR   GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR   GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 2.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Cell adhesion; Integrin; Transmembrane.
SQ   SEQUENCE   1151 AA;  126943 MW;  8F785695D4074CA5 CRC64;

  Query Match             74.6%;  Score 4381;  DB 2;  Length 1151;
  Best Local Similarity   72.7%;  Pred. No. 6.2e-281;
  Matches  827;  Conservative  150;  Mismatches  158;  Indels    2;  Gaps    2;
```

```
Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        796 DSRDFDVSVTLRNDGEDSYGTKVTCYYPSGLSYRKVSASQNQFSKKPWRVIAE-PSSSEG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 QGVLKSTIWDINHPIFPANSEVTFNVTFDVDSDASLGNKLLLKVVVTSENNVSRTDKTES 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        915 QLELPVKYAIYMVVTSGESSIKYLNFTASEMTSKVIQHQYQFNNLGQRSFPVSVVFWIPV 974

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        975 QINKVTIWDPPQVTFSQNLSSVCRTEQKSPSHSKFQDELERTPVLNCSVAVCKRIQCDIP 1034

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1035 SFNSKEIFNVTLQGNLLFDWYIETSHDHLLLVSTAEILFNDSAFALLPGQETFVKAQTET 1094

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1095 KVEPYTVHNPVPLIVGSSVGGLVLLALITAGLYKLGFFKRQYKDMMNEAGGQDGPPQ 1151
```

RESULT 4
Q28984
ID Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CD11b (Fragment).
GN Name=CD11b;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; U40072; AAB16869.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Cell adhesion; Integrin; Transmembrane.
FT NON_TER 1 1
FT NON_TER 920 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;

Query Match 64.8%; Score 3809.5; DB 2; Length 920;
Best Local Similarity 78.9%; Pred. No. 3.2e-243;
Matches 727; Conservative 84; Mismatches 109; Indels 1; Gaps 1;

```
Qy        118 PQKFPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYS 177
Db          1 PQKFPEALRGCPQQESDIAFLIDGSGSINRLDFQRMKEFVSTVMGQFQKSKTLFALMQYS 60

Qy        178 EEFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKIL 237
Db         61 EDFYTHFTFNDFKRNPSPKLLVRPIRQLLGRTHTATGIRKVVRELFHSKSGARENALKIL 120

Qy        238 ILITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVF 297
Db        121 VVITDGEKFGDPLGYEDVIPEADRKGVIRYVIGVGDAFNSWKSREELNTIASKPCGDHVF 180

Qy        298 QINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYD 357
Db        181 QVNNFEAVKTIQNQLQEKTFAIEGTQTGSTSSFECEMSQEGFSAAITSNGPLLGAVGSFD 240

Qy        358 WAGGVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVA 417
Db        241 WAGGAFLHMPKDRVIFINTTRVDSDMNDAYLGYAVEVILRNQAQSLVLGAPRYQHTGLVV 300

Qy        418 MFRQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSV 477
Db        301 MFKQNSGAWEKNADIKGSQIGSYFGASLCSVDVNRDGSSDLVLIGAPHYYEQTRGGQVSV 360

Qy        478 CPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAV 537
Db        361 CPLPQG-RAKWQCRVILCGEQGHPWSRFGAALTALGDVNGDKLTDVAIGAPGEQDNRGAV 419

Qy        538 YLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLL 597
Db        420 YLFHGTSELGISPSHSQRIAGSQLSPRLQYFGQSLSGGQDLTMDGLMDLAVGAQGHVFLL 479

Qy        598 RSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQS 657
Db        480 RSQPVLRVEASMVFKPREVARNVYECRQQAAKTQIAGEVQVCLQVRKSTWDRLREGDTQS 539

Qy        658 VVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLR 717
Db        540 IITYDLALDPGRPHPRAVFEETKNNTRRQTQTLGLSRKCEHLALWLPDCVEDSVTPIVLR 599

Qy        718 LNFSLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCL 777
Db        600 LNFSLVGKPASSFGNLRPVLAVDAQRLFTALFPPEKNCGNDSICQDDLSITFSFMSLDTL 659

Qy        778 VVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASS 837
Db        660 VVGGPRDLKVTLTVRNQGEDSYRTQVTFFYPSDLSYRKVSTSQNQRSQRSWRLACESDVS 719

Qy        838 TEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNK 897
Db        720 TEESTALKSTSCSINHPIFPDNSEVTFNVTFDVNPDAFLGYKLLLKANVTSENNMPSSNK 779

Qy        898 TEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFL 957
Db        780 TEFQLELPVKYAVYVVVTSLEVSTKYFNFTASEKTRHVIEHQYQFNNLGQRKLPISVVFW 839

Qy        958 VPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQC 1017
Db        840 VPVRLNRVTVWDQPQVTFSQNLSRSCSTEEIGPRHSDFLEKLQKTPVLNCSIAVCQKIQC 899

Qy       1018 DIPFFGIQEEFNATLKGNLSF 1038
Db        900 DIPSFGIQEELKVTLKGNLSF 920
```

RESULT 5
ITAX_HUMAN
ID ITAX_HUMAN STANDARD; PRT; 1163 AA.
AC P20702; Q8IVA6;
DT 01-FEB-1991 (Rel. 17, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
GN Name=ITGAX; Synonyms=CD11C;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```
RX   MEDLINE=88166645; PubMed=3327687;
RA   Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT   "cDNA cloning and complete primary structure of the alpha subunit of a
RT   leukocyte adhesion glycoprotein, p150,95.";
RL   EMBO J. 6:4023-4028(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90153906; PubMed=2303426;
RA   Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT   "Genomic structure of an integrin alpha subunit, the leukocyte p150,95
RT   molecule.";
RL   J. Biol. Chem. 265:2782-2788(1990).
RN   [3]
RP   ERRATUM.
RA   Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL   J. Biol. Chem. 265:12750-12751(1990).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 20-43.
RX   MEDLINE=87167596; PubMed=3549901;
RA   Miller L.J., Wiebe M., Springer T.A.;
RT   "Purification and alpha subunit N-terminal sequences of human Mac-1
RT   and p150,95 leukocyte adhesion proteins.";
RL   J. Immunol. 138:2381-2383(1987).
CC   -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
CC       recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
CC       interaction during inflammatory responses. It is especially
CC       important in monocyte adhesion and chemotaxis.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC       granulocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M81695; AAA59180.1; -.
DR   EMBL; M29165; -; NOT_ANNOTATED_CDS.
DR   EMBL; M29487; AAA51620.1; ALT_SEQ.
DR   EMBL; M29482; AAA51620.1; JOINED.
DR   EMBL; M29483; AAA51620.1; JOINED.
DR   EMBL; M29484; AAA51620.1; JOINED.
DR   EMBL; M29485; AAA51620.1; JOINED.
DR   EMBL; M29486; AAA51620.1; JOINED.
DR   EMBL; BC038237; AAH38237.1; -.
DR   PIR; A36584; RWHU1C.
DR   PDB; 1N3Y; X-ray; A=141-338.
DR   Genew; HGNC:6152; ITGAX.
DR   MIM; 151510; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0004872; F:receptor activity; TAS.
DR   GO; GO:0007155; P:cell adhesion; TAS.
DR   GO; GO:0009887; P:organogenesis; TAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
KW   Glycoprotein; Integrin; Magnesium; Polymorphism; Receptor; Repeat;
KW   Signal; Transmembrane.
FT   SIGNAL       1     19
FT   CHAIN       20   1163       Integrin alpha-X.
FT   DOMAIN      20   1107       Extracellular (Potential).
FT   TRANSMEM  1108   1128       Potential.
FT   DOMAIN    1129   1163       Cytoplasmic (Potential).
FT   REPEAT      34     87       FG-GAP 1.
FT   REPEAT      88      ?       FG-GAP 2.
FT   DOMAIN     165    351       VWFA.
FT   REPEAT       ?    401       FG-GAP 3.
FT   REPEAT     402    453       FG-GAP 4.
FT   REPEAT     455    517       FG-GAP 5.
FT   REPEAT     518    576       FG-GAP 6.
FT   REPEAT     581    633       FG-GAP 7.
FT   CA_BIND    466    474       Potential.
FT   CA_BIND    530    538       Potential.
FT   CA_BIND    593    601       Potential.
FT   SITE      1131   1135       GFFKR motif.
FT   DISULFID    69     76       By similarity.
FT   DISULFID   108    126       By similarity.
FT   DISULFID   655    712       By similarity.
FT   DISULFID   771    777       By similarity.
FT   DISULFID   848    863       By similarity.
FT   DISULFID   998   1022       By similarity.
FT   DISULFID  1027   1032       By similarity.
FT   CARBOHYD    61     61       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   392    392       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   697    697       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   735    735       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   899    899       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   939    939       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1050   1050       N-linked (GlcNAc. . .) (Potential).
FT   VARIANT     48     48       W -> R (in dbSNP:11574633).
FT                               /FTId=VAR_018672.
FT   CONFLICT   209    209       T -> S (in Ref. 4).
FT   CONFLICT   251    251       T -> A (in Ref. 4).
FT   CONFLICT   469    469       T -> S (in Ref. 4).
FT   CONFLICT   490    490       G -> A (in Ref. 2).
FT   CONFLICT   547    547       E -> K (in Ref. 4).
FT   CONFLICT   756    756       D -> L (in Ref. 1).
FT   CONFLICT   819    819       I -> V (in Ref. 4).
FT   CONFLICT  1161   1163       SEK -> TPHYPQDNV (in Ref. 4).
FT   STRAND     150    157
```

```
FT   TURN        160    161
FT   HELIX       164    178
FT   TURN        179    180
FT   TURN        183    185
FT   STRAND      186    193
FT   STRAND      197    201
FT   HELIX       203    208
FT   HELIX       212    216
FT   TURN        217    218
FT   STRAND      226    226
FT   HELIX       228    236
FT   TURN        237    240
FT   HELIX       242    244
FT   TURN        245    245
FT   TURN        248    249
FT   STRAND      251    258
FT   STRAND      263    263
FT   HELIX       269    278
FT   TURN        279    280
FT   STRAND      282    288
FT   HELIX       290    293
FT   TURN        296    297
FT   HELIX       298    304
FT   HELIX       310    312
FT   STRAND      313    316
FT   HELIX       319    325
FT   HELIX       326    334
FT   TURN        335    335
SQ   SEQUENCE   1163 AA;  127886 MW;  83658A13B5C5DE8F CRC64;

  Query Match             59.2%;  Score 3479;  DB 1;  Length 1163;
  Best Local Similarity   61.1%;  Pred. No. 3.6e-221;
  Matches  690;  Conservative  141;  Mismatches  292;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||| |: :: ||| ||||  | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||:|||| ||||||||||  |  | |: |||||  |    |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEPISTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || ||:||||||   :|      |:  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              : ||||:||:   || ||:  : || | |:||| |: ||  ||: : |||::| ||||:|
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:| || | |:|||| ||  |:||| |||| ||::  | :||| :|||| ::|:|::
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:\|| |||||:|||||||||:| |||||| ||:|||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |  ||||||||||||| |    :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:|||||||||||::|||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||||| |||||||:|||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || |  |||||||||||:|| |||||||:|||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |  |:| | |: |: ||| :|||  :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || |||||:|| |||| ||| |||||||| |:||||:| |:||| |  |:||
Db        737 TLVGKPLLAFRNLRPMLAADAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |  ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::||  |||| || ||:::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| ||:||  ||||||: | |  | |  |:|||:||||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|  :|:  :| |  | :::  |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :      : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143


RESULT 6
ITAD_HUMAN
ID   ITAD_HUMAN     STANDARD;      PRT;  1162 AA.
AC   Q13349; Q15575; Q15576;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN   Name=ITGAD;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=96111956; PubMed=8777714;
RA   Van der Vieren M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA   Staunton D.E., Gallatin W.M.;
RT   "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
RT   3.";
RL   Immunity 3:683-690(1995).
RN   [2]
RP   SEQUENCE OF 1-235 FROM N.A.
RX   MEDLINE=20187620; PubMed=10722744;
RA   Noti J.D., Johnson A.K., Dillon J.D.;
RT   "Structural and functional characterization of the leukocyte integrin
RT   gene CD11d. Essential role of Sp1 and Sp3.";
RL   J. Biol. Chem. 275:8959-8969(2000).
RN   [3]
RP   SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX   MEDLINE=96257236; PubMed=8666289;
RA   Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
RT   "Cloning and chromosomal localization of a novel gene-encoding a human
RT   beta 2-integrin alpha subunit.";
```

```
RL   Gene 171:291-294(1996).
RN   [4]
RP   INTERACTION WITH VCAM1.
RX   MEDLINE=99059842; PubMed=9841932;
RA   Grayson M.H., Van der Vieren M., Sterbinsky S.A., Michael Gallatin W.,
RA   Hoffman P.A., Staunton D.E., Bochner B.S.;
RT   "alphadbeta2 integrin is expressed on human eosinophils and functions
RT   as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-
RT   1).";
RL   J. Exp. Med. 188:2187-2191(1998).
RN   [5]
RP   INTERACTION WITH VCAM1.
RX   MEDLINE=99370002; PubMed=10438935;
RA   Van der Vieren M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
RA   Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RT   "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
RT   binding interface between I domain and VCAM-1.";
RL   J. Immunol. 163:1984-1990(1999).
CC   -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC       VCAM1. May play a role in the atherosclerotic process such as
CC       clearing lipoproteins from plaques and in phagocytosis of blood-
CC       borne pathogens, particulate matter, and senescent erythrocytes
CC       from the blood.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell
CC       lines and subsets of peripheral blood leukocytes and strongly on
CC       tissue-specialized cells, including macrophages foam cells within
CC       atherosclerotic plaques, and on splenic red pulp macrophages.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U37028; AAB38547.1; -.
DR   EMBL; U40274; AAB60634.1; -.
DR   EMBL; U40275; AAB60635.1; -.
DR   EMBL; U40276; AAB60636.1; -.
DR   EMBL; U40277; AAB60637.1; -.
DR   EMBL; U40279; AAB60638.1; -.
DR   EMBL; U40278; AAB60638.1; JOINED.
DR   EMBL; AF187881; AAF62875.1; -.
DR   HSSP; P11215; 1BHQ.
DR   Genew; HGNC:6146; ITGAD.
DR   MIM; 602453; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0016337; P:cell-cell adhesion; NAS.
DR   GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     17       Potential.
FT   CHAIN        18   1162       Integrin alpha-D.
FT   DOMAIN       18   1100       Extracellular (Potential).
FT   TRANSMEM   1101   1121       Potential.
FT   DOMAIN     1122   1162       Cytoplasmic (Potential).
FT   REPEAT       32     85       FG-GAP 1.
FT   REPEAT       86      ?       FG-GAP 2.
FT   DOMAIN      150    332       VWFA.
FT   REPEAT      350    400       FG-GAP 3.
FT   REPEAT      401    452       FG-GAP 4.
FT   REPEAT      454    516       FG-GAP 5.
FT   REPEAT      518    576       FG-GAP 6.
FT   REPEAT      581    633       FG-GAP 7.
FT   CA_BIND     465    473       Potential.
FT   CA_BIND     530    538       Potential.
FT   CA_BIND     593    601       Potential.
FT   SITE       1127   1131       GFFKR motif.
FT   DISULFID     67     74       By similarity.
FT   DISULFID    106    124       By similarity.
FT   DISULFID    655    710       By similarity.
FT   DISULFID    769    775       By similarity.
FT   DISULFID    846    861       By similarity.
FT   DISULFID    994   1018       By similarity.
FT   DISULFID   1023   1028       By similarity.
FT   CARBOHYD     59     59       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     87     87       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     99     99       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    391    391       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    691    691       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    733    733       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    873    873       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    957    957       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1046   1046       N-linked (GlcNAc. . .) (Potential).
FT   CONFLICT    500    500       Missing (in Ref. 2).
FT   CONFLICT    515    518       GHPW -> ATP (in Ref. 2).
FT   CONFLICT    825    825       L -> V (in Ref. 2).
FT   CONFLICT    984    984       V -> A (in Ref. 2).
SQ   SEQUENCE   1162 AA;  126885 MW;  F296A1A35455D77D CRC64;

  Query Match             58.0%;  Score 3411;  DB 1;  Length 1162;
  Best Local Similarity   59.4%;  Pred. No. 1.1e-216;
  Matches  671;  Conservative  166;  Mismatches  284;  Indels    8;  Gaps    6;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    |||:| |||||||| |||:||||| |:||||| | || | :|| |:||
Db         18 FNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQTGRLYDCAAATGMCQPI 77

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              | :  |||||||||:|||:|:  :|||||||:|: | ||:| || | | ||   :  |
Db         78 PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-EIIQT 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
               |:|   || :: || ||:||||||  :|| : | |:  || | : : |||:|||||
Db        137 VPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFEGTDTLFALMQYSNLL 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              :||||| :|: :|: :||: || || | | |||||  || :||:  |||||:| ||||:|
Db        197 KIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHHKNGARKSAKKILIVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:|: ||| | ||||:|:: |:||| |||| ||:   :||||||::| ||:||||:::
Db        257 TDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELNTISSAPPQDHVFKVD 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              || || :|| ||:|||:|:|||| :||||:|||||||| |:| :|  |  |||: |:|
Db        317 NFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGSFSWSG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||       |||||:: : || |:||||:  : |   ||:|||||||||| |   :|
Db        377 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVLGAPRYQHTGKAVIFT 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:||||||||||||:|||||:|||||||||||||||||||||
Db        437 QVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPHYYEQTRGGQVSVCPL 496
```

```
Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||| :|||||||| |||| ||||||||||||||| ||| |||||||||::||||||||
Db        497 PRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAIGAPGEQENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || | |||||||||||| |:|||||||||:|||||||| |||:|| |||:| ||||||
Db        557 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLLRSL 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |||:|   | |:| |||: |: | ::     |||:  ||| :|||: |:|  | ||| |
Db        617 PVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 674

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              :||||| ||  |||:|||||| |  : : |||   |||||| ||:|:|| ||||:| |||
Db        675 FDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCVEDVVSPIILHLNF 734

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||  |: :  |||||||  :| ||||  |||||||  | :|: || :| ||  |  | ||
Db        735 SLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDLGVTLSFSGLQTLTVG 794

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | || ||| | ||||| | |: ::|  ||:|:||  | |  | : |||||:   ||
Db        795 SSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 853

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              | |:|: ||:||||| | |  || :|||  ||:||:::|::|: :|||   ::|  |
Db        854 EG-LRSSRCSVNHPIFHEGSNGTFIVTFDVSYKATLGDRMLMRASASSENNKASSSKATF 912

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNF-TASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||||| :::    |||| || |: |   :  :|:|:|:|:|| || | ||: | ||
Db        913 QLELPVKYAVYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQRDLAISINFWVP 972

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | ||   :||      |::|   | :: : | ||||| :: ::|:::|||| | : :||:
Db        973 VLLNGVAVWDVVMEAPSQSL--PCVSERKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1030

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| | :|    :|:|| ||| |: ||:: |||| ||:|:| |
Db       1031 PSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQLPGQEAFMRAQME 1090

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
                :|  || | :|:|:||||| ||||||||| |||||||| ||:|: :
Db       1091 MVLEEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRHYKEMLED 1139
```

```
RESULT 7
Q6KAS4
ID   Q6KAS4      PRELIMINARY;      PRT;  1188 AA.
AC   Q6KAS4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MFLJ00114 protein (Fragment).
GN   Name=mFLJ00114;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA   Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT   "Prediction of the Coding Sequences of Mouse Homologues of FLJ Genes:
RT   The Complete Nucleotide Sequences of 110 Mouse FLJ-Homologous cDNAs
RT   Identified by Screening of Terminal Sequences of cDNA Clones Randomly
RT   Sampled from Size-Fractionated Libraries.";
RL   DNA Res. 11:167-180(2004).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR   EMBL; AK131133; BAD21383.1; -.
DR   GO; GO:0009897; C:external side of plasma membrane; IDA.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; UNKNOWN_1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Cell adhesion; Integrin; Transmembrane.
FT   NON_TER      1      1
SQ   SEQUENCE   1188 AA;  131248 MW;  B8D93107BDBB4178 CRC64;

  Query Match             56.7%;  Score 3329.5;  DB 2;  Length 1188;
  Best Local Similarity   56.9%;  Pred. No. 3e-211;
  Matches  649;  Conservative  170;  Mismatches  302;  Indels   19;  Gaps    7;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    | :   || ||:|   | |||||||:|| | || | ||:| | ||:||||
Db         39 FNLDAEKPTHFHMDGAEFGHSVLQYDSSWVVVGAPKEIKATNQIGGLYKCGYHTGNCEPI 98

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||||| |:|  |||||||| || || |: ||||| |: :| |
Db         99 SLQVPPEAVNMSLGLSLAAATNPSWLLACGPTVHHTCRENIYLTGLCFLLSSSFKQS-QN 157

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              || | : ||::| || ||:||||||   || :  :|:  || ||::  | |||||:|: |
Db        158 FPTAQQECPKQDQDIVFLIDGSGSISSTDFEKMLDFVKAVMSQLQRPSTRFSLMQFSDYF 217

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              |:|||  | :  :| ||:  :  || | |:||: |: |: |||   :|||::|  |:||:|
Db        218 RVHFTFNNFISTSSPLSLLDSVRQLRGYTYTASAIKHVITELFTTQSGARQDATKVLIVI 277

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||| | || | |: ||| |:   :||| |||| || :| |:||| :|| |  ::|| :
Db        278 TDGRKQGDNLSYDSVIPMAEAASIIRYAIGVGKAFYNEHSKQELKAIASMPSHEYVFSVE 337

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||:||| |:|||:|||||||||:| |||:|| |||||||| | :||:|  |||: |:|
Db        338 NFDALKDIENQLKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTPDGPVLGAVGSFSWSG 397

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | ||| |  : |||||:: : |||||: |:      | ||:|||||:|| | | :|
Db        398 GAFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVVIFT 457

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   : |:|||||:|||||||||:| :|||||||| ||||| ||||||||||:
Db        458 QESRHWRPKSEVRGTQIGSYFGASLCSVDMDRDGSTDLVLIGVPHYYEHTRGGQVSVCPM 517

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              | |   || |   |:|||| |||||||||||||||| | ||||||||||:||||||:|
Db        518 P-GVGRRWHCGTTLHGEQGHPWGRFGAALTVLGDVNGDSLADVAIGAPGEEENRGAVYIF 576

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || |   |:|| |||| |::  |:|||||||||||||| |||||  |||::| |||||::
Db        577 HGASRQDIAPSPSQRISASQIPSRIQYFGQSLSGGQDLTRDGLVDLAVGSKGRVLLLRTR 636

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |:|||   : | | |::|:|||| :||   :   :  ||||: :| : :|  | ::| ||
Db        637 PILRVSPTVHFTPAEISRSVFECQEQVAPEQTLSDATVCLHIHESPKTQL--GDLRSTVT 694

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              :||||| ||  :||:| |||     : : ||| : ||::|| || |:|| |:|| |||||
Db        695 FDLALDHGRLSTRAIFKETKTRALTRVKTLGLNKHCESVKLLLPACVEDSVTPITLRLNF 754

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||| |:|:  ||:|:||| | |  |||  |||||||  |:|||||||: | |  |  ||||
Db        755 SLVGVPISSLQNLQPMLAVDDQTYFTASLPFEKNCGADHICQDDLSVVFGFPDLKTLVVG 814

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQ----------NQRSQRSWR 829
              | || ||| ||||||| | || |:|: ||:|:|:  |            :| | |
Db        815 SDLELNVDVTVSNDGEDSYGTTVTLFYPVGLSFRRVAEGQVFLRKKEDQQWQRRGQHSLH 874
```

```
Qy        830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSE 889
Db        875 LMCD--STPDRSQGLWSTSCSSRHVIFRGGSQMTFLVTFDVSPKAELGDRLLLRARVGSE 932

Qy        890 NNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASE-NTSRVMQHQYQVSNLGQR 948
Db        933 NNVPGTPKTTFQLELPVKYAVYTMISSHDQFTKYLNFSTSEKEKTSVVEHRFQVNNLGQR 992

Qy        949 SLPISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCS 1008
Db        993 DVPVSINFWVPIELKGEAVW-TVMVSHPQNPLTQCYRNRLKPTQFDFLTHMQKSPVLDCS 1051

Qy       1009 IAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLP 1068
Db       1052 IADCLHLRCDIPSLGILDELYFILKGNLSFGWISQTLQKKVLLLSEAEITFNTSVYSQLP 1111

Qy       1069 GQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1112 GQEAFLRAQTKTVLEMYKVHNPVPLIVGSSVGGLLLLAIITAILYKAGFFKRQYKEMLEE 1171
```

RESULT 8
BAD21383
ID BAD21383 PRELIMINARY; PRT; 1188 AA.
AC BAD21383;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE MFLJ00114 protein (Fragment).
GN MFLJ00114.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Gene:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-homologous cDNAs
RT Identified by Screening of Terminal sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK131133; BAD21383.1; -.
FT NON_TER 1 1
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDBB4178 CRC64;

Query Match 56.7%; Score 3329.5; DB 2; Length 1188;
Best Local Similarity 56.9%; Pred. No. 3e-211;
Matches 649; Conservative 170; Mismatches 302; Indels 19; Gaps 7;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         39 FNLDAEKPTHFHMDGAEFGHSVLQYDSSWVVVGAPKEIKATNQIGGLYKCGYHTGNCEPI 98

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         99 SLQVPPEAVNMSLGLSLAAATNPSWLLACGPTVHHTCRENIYLTGLCFLLSSSFKQS-QN 157

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
Db        158 FPTAQQECPKQDQDIVFLIDGSGSISSTDFEKMLDFVKAVMSQLQRPSTRFSLMQFSDYF 217

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
Db        218 RVHFTFNNFISTSSPLSLLDSVRQLRGYTYTASAIKHVITELFTTQSGARQDATKVLIVI 277

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db        278 TDGRKQGDNLSYDSVIPMAEAASIIRYAIGVGKAFYNEHSKQELKAIASMPSHEYVFSVE 337

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        338 NFDALKDIENQLKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTPDGPVLGAVGSFSWSG 397

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        398 GAFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVVIFT 457

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        458 QESRHWRPKSEVRGTQIGSYFGASLCSVDMDRDGSTDLVLIGVPHYYEHTRGGQVSVCPM 517

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        518 P-GVGRRWHCGTTLHGEQGHPWGRFGAALTVLGDVNGDSLADVAIGAPGEEENRGAVYIF 576

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        577 HGASRQDIAPSPSQRISASQIPSRIQYFGQSLSGGQDLTRDGLVDLAVGSKGRVLLLRTR 636

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        637 PILRVSPTVHFTPAEISRSVFECQEQVAPEQTLSDATVCLHIHESPKTQL--GDLRSTVT 694

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        695 FDLALDHGRLSTRAIFKETKTRALTRVKTLGLNKHCESVKLLLPACVEDSVTPITLRLNF 754

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        755 SLVGVPISSLQNLQPMLAVDDQTYFTASLPFEKNCGADHICQDDLSVVFGFPDLKTLVVG 814

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQ-----------NQRSQRSWR 829
Db        815 SDLELNVDVTVSNDGEDSYGTTVTLFYPVGLSFRRVAEGQVFLRKKEDQQWQRRGQHSLH 874

Qy        830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSE 889
Db        875 LMCD--STPDRSQGLWSTSCSSRHVIFRGGSQMTFLVTFDVSPKAELGDRLLLRARVGSE 932

Qy        890 NNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASE-NTSRVMQHQYQVSNLGQR 948
Db        933 NNVPGTPKTTFQLELPVKYAVYTMISSHDQFTKYLNFSTSEKEKTSVVEHRFQVNNLGQR 992

Qy        949 SLPISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCS 1008
Db        993 DVPVSINFWVPIELKGEAVW-TVMVSHPQNPLTQCYRNRLKPTQFDFLTHMQKSPVLDCS 1051

Qy       1009 IAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLP 1068
Db       1052 IADCLHLRCDIPSLGILDELYFILKGNLSFGWISQTLQKKVLLLSEAEITFNTSVYSQLP 1111

Qy       1069 GQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1112 GQEAFLRAQTKTVLEMYKVHNPVPLIVGSSVGGLLLLAIITAILYKAGFFKRQYKEMLEE 1171
```

RESULT 9
ITAX_MOUSE
ID ITAX_MOUSE STANDARD; PRT; 1169 AA.
AC Q9QXH4;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c).
GN Name=Itgax;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Dendritic cell;

```
RA   Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RA   Tsuchiya H.;
RT   "Isolation of genes selectively expressed by dendritic cells.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
CC       recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
CC       interaction during inflammatory responses. It is especially
CC       important in monocyte adhesion and chemotaxis (By similarity).
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
CC       associates with beta-2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF211864; AAF23492.1; -.
DR   HSSP; P20702; 1N3Y.
DR   MGD; MGI:96609; Itgax.
DR   GO; GO:0009897; C:external side of plasma membrane; IDA.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL       1     19       Potential.
FT   CHAIN       20   1169       Integrin alpha-X.
FT   DOMAIN      20   1116       Extracellular (Potential).
FT   TRANSMEM  1117   1137       Potential.
FT   DOMAIN    1138   1169       Cytoplasmic (Potential).
FT   REPEAT      34     87       FG-GAP 1.
FT   REPEAT      88      ?       FG-GAP 2.
FT   DOMAIN     152    330       VWFA.
FT   REPEAT       ?    402       FG-GAP 3.
FT   REPEAT     403    454       FG-GAP 4.
FT   REPEAT     456    518       FG-GAP 5.
FT   REPEAT     519    577       FG-GAP 6.
FT   REPEAT     582    634       FG-GAP 7.
FT   CA_BIND    467    475       Potential.
FT   CA_BIND    531    539       Potential.
FT   CA_BIND    594    602       Potential.
FT   SITE      1140   1144       GFFKR motif.
FT   DISULFID    69     76       By similarity.
FT   DISULFID   108    126       By similarity.
FT   DISULFID   656    711       By similarity.
FT   DISULFID   770    776       By similarity.
FT   DISULFID   858    873       By similarity.
FT   DISULFID  1007   1031       By similarity.
FT   DISULFID  1036   1041       By similarity.
FT   CARBOHYD    89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   267    267       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   393    393       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   734    734       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   949    949       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1059   1059       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1084   1084       N-linked (GlcNAc. . .) (Potential).
```

```
SQ   SEQUENCE   1169 AA;  129150 MW;  C616412033C219A6 CRC64;

  Query Match             56.5%;  Score 3320.5;  DB 1;  Length 1169;
  Best Local Similarity   56.8%;  Pred. No. 1.1e-210;
  Matches  647;  Conservative  172;  Mismatches  302;  Indels   19;  Gaps    7;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |     |  :   || ||:|    | ||||||:|| | || | ||:| | ||:|||
Db         20 FNLDAEKLTHFHMDGAEFGHSVLQYDSSWVVVGAPKEIKATNQIGGLYKCGYHTGNCEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||:|||||||| |:|  |||||||| || || |: ||||  |: :|  |
Db         80 SLQVPPEAVNISLGLSLAAATNPSWLLACGPTVHHTCRENIYLTGLCFLLSSSFKQS-QN 138

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              || | : ||::| || ||:||||||   || :  :|:  || ||::  | |||||:|: |
Db        139 FPTAQQECPKQDQDIVFLIDGSGSISSTDFEKMLDFVKAVMSQLQRPSTRFSLMQFSDYF 198

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              |:|||  | :  :| ||:  : || | |:||: |: |: |||   :|||::| |:||:|
Db        199 RVHFTFNNFISTSSPLSLLGSVRQLRGYTYTASAIKHVITELFTTQSGARQDATKVLIVI 258

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||| | || | |: ||| |:   :||| |||| || :| |:|||  :|| |  ::|| :
Db        259 TDGRKQGDNLSYDSVIPMAEAASIIRYAIGVGKAFYNEHSKQELKAIASMPSHEYVFSVE 318

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||:||| |:|||:|||||||||:| |||:|| |||||||||  | :||:|  |||: |:|
Db        319 NFDALKDIENQLKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTPDGPVLGAVGSFSWSG 378

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | ||| |  : |||||:: : || |||||: |:     | ||:|||||:|| | | :|
Db        379 GAFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVVIFT 438

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   : |:|||||:||||||||||:| :|||||||| ||||| ||||||||||:
Db        439 QESRHWRPKSEVRGTQIGSYFGASLCSVDMDRDGSTDLVLIGVPHYYEHTRGGQVSVCPM 498

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              | | :|| |   |:|||| |||||||||||||||||| | ||||||||||:||||||:|
Db        499 P-GVGSRWHCGTTLHGEQGHPWGRFGAALTVLGDVNGDSLADVAIGAPGEEENRGAVYIF 557

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || |   |:|| |||| |::  |:|||||||||||| |||||  ||::| |||||:
Db        558 HGASRQDIAPSPSQRISASQIPSRIQYFGQSLSGGQDLTRDGLVDLAVGSKGRVLLLRTR 617

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |:|||   : | | |::|:|||| :||   :   :  ||||: :| : :|  | ::| ||
Db        618 PILRVSPTVHFTPAEISRSVFECQEQVAPEQTLSDATVCLHIHESPKTQL--GDLRSTVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              :||||| ||  :||:| |||     : : ||| : ||::|| || |:|| |:|| ||||
Db        676 FDLALDHGRLSTRAIFKETKTRALTRVKTLGLNKHCESVKLLLPACVEDSVTPITLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||| |:|:  ||:|:|| | |  |||  ||||||| |:|||||: | |  |  ||||
Db        736 SLVGVPISSLQNLQPMLAVDDQTYFTASLPFEKNCGADHICQDDLSVVFGFPDLKTLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQ----------NQRSQRSWR 829
              | || ||| ||||||| | || |:|: ||:|:|:  |            :| | |
Db        796 SDLELNVDVTVSNDGEDSYGTTVTLFYPVGLSFRRVAEGQVFLRKKEDQQWQRRGQHSLH 855

Qy        830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSE 889
              | |:  |: : |  | |||||  | ||   |::|| :|||| || ||::|||:| | ||
Db        856 LMCD--STPDRSQGLWSTSCSSRHVIFRGGSQMTFLVTFDVSPKAELGDRLLLRARVGSE 913

Qy        890 NNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASE-NTSRVMQHQYQVSNLGQR 948
              ||:| | || |||||||||||| :::||   |||||: ||   : |::|::|:|||||
Db        914 NNVPGTPKTTFQLELPVKYAVYTMISSHDQFTKYLNFSTSEKEKTSVVEHRFQVNNLGQR 973

Qy        949 SLPISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCS 1008
              :|:|: | ||: |    :|    |:  :|  : |:      |:  | |  ::|:||::||
Db        974 DVPVSINFWVPIELKGEAVW-TVMVSHPQNPLTQCYRNRLKPTQFDLLTHMQKSPVLDCS 1032
```

```
Qy       1009 IAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLP 1068
              || |  ::|||| || :|    |||||| |  :|    :|::| ||| || ||:: ||
Db       1033 IADCLHLRCDIPSLGILDELYFILKGNLSFGWISQTLQKKVLLLSEAEITFNTSVYSQLP 1092

Qy       1069 GQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              || ||:|:||:| :| ::| ||:||||||||||||||:||| ||| ||||||||:|: |
Db       1093 GQEAFLRAQTKTVLEMYKVHNPVPLIVGSSVGGLLLLAIITAILYKAGFFKRQYKEMLEE 1152


RESULT 10
ITAD_RAT
ID   ITAD_RAT       STANDARD;      PRT;  1161 AA.
AC   Q9QYE7;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-D precursor.
GN   Name=Itgad;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RA   O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA   Gallatin W.M.;
RT   "Cloning of rat alpha D, a novel beta 2 integrin.";
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC       VCAM1. May play a role in the atherosclerotic process such as
CC       clearing lipoproteins from plaques and in phagocytosis of blood-
CC       borne pathogens, particulate matter, and senescent erythrocytes
CC       from the blood (By similarity).
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC       associates with beta-2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF021334; AAF21241.1; -.
DR   HSSP; P11215; 1BHQ.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     19       Potential.
FT   CHAIN        20   1161       Integrin alpha-D.
FT   DOMAIN       20   1100       Extracellular (Potential).
FT   TRANSMEM   1101   1121       Potential.
FT   DOMAIN     1122   1161       Cytoplasmic (Potential).
FT   REPEAT       34     87       FG-GAP 1.
FT   REPEAT       88      ?       FG-GAP 2.
FT   DOMAIN      152    334       VWFA.
FT   REPEAT      352    402       FG-GAP 3.
FT   REPEAT      403    454       FG-GAP 4.
FT   REPEAT      456    517       FG-GAP 5.
FT   REPEAT      519    577       FG-GAP 6.
FT   REPEAT      582    634       FG-GAP 7.
FT   CA_BIND     467    475       Potential.
FT   CA_BIND     531    539       Potential.
FT   CA_BIND     594    602       Potential.
FT   SITE       1126   1130       GFFKR motif.
FT   DISULFID     69     76       By similarity.
FT   DISULFID    108    126       By similarity.
FT   DISULFID    656    711       By similarity.
FT   DISULFID    769    775       By similarity.
FT   DISULFID    845    860       By similarity.
FT   DISULFID    993   1017       By similarity.
FT   DISULFID   1022   1027       By similarity.
FT   CARBOHYD     61     61       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    245    245       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    393    393       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    696    696       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    734    734       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    784    784       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    907    907       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    936    936       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1045   1045       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   1161 AA;  126600 MW;  2258491A984A705E CRC64;

  Query Match             55.2%;  Score 3244.5;  DB 1;  Length 1161;
  Best Local Similarity   57.6%;  Pred. No. 1.2e-205;
  Matches  649;  Conservative  163;  Mismatches  302;  Indels   13;  Gaps    9;

Qy          2 NLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPIR 61
              ||| |  : |:|:|  |||:||| |||:||||| | || || | || |  :|| |:||
Db         21 NLDVEEPIVFREDAASFGQTVVQFGGSRLVVGAPLEAVAVNQTGRLYDCAPATGMCQPIV 80

Qy         62 LQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQKF 121
              |: |:||||||||||||   |:  |||||||  : | :| | || | | ||:| |  |
Db         81 LRSPLEAVNMSLGLSLVTATNNAQLLACGPTAQRACVKNMYAKGSCLLLGSSL-QFIQAV 139

Qy        122 PEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEFR 181
              | ::  ||::: |||||:|||||   || : |:|:  :| :   : |||||||||   :
Db        140 PASMPECPRQEMDIAFLIDGSGSINQRDFAQMKDFVKALMGEFASTSTLFSLMQYSNILK 199

Qy        182 IHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILIT 241
              |||| ||:|  :|:||: || || | |:|||||| |: |||:  ||:||:| |||::||
Db        200 THFTFTEFKNILDPQSLVDPIVQLQGLTYTATGIRTVMEELFHSKNGSRKSAKKILLVIT 259

Qy        242 DGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQINN 301
              ||:|: ||| | |||| ||: |:||| |||||||:   : :||||: | ||:||||:: |
Db        260 DGQKYRDPLEYSDVIPAADKAGIIRYAIGVGDAFQEPTALKELNTIGSAPPQDHVFKVGN 319

Qy        302 FEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAGG 361
              | ||::|| ||:||||||||||: |||||:|||||||||:|:||:||:|  |||: |:||
Db        320 FAALRSIQRQLQEKIFAIEGTQSRSSSSFQHEMSQEGFSSALTSDGPVLGAVGSFSWSGG 379

Qy        362 VFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFRQ 421
              |||    : |||||:: : || |:||||: |:     | ||:|||||:|| | | :| |
Db        380 AFLYPPNTRPTFINMSQENVDMRDSYLGYSTAVAFWKGVHSLILGAPRHQHTGKVVIFTQ 439

Qy        422 NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPLP 481
              |   : |:|||||:|||||||||||| :|||||||||||||||||||||||||| |:|
Db        440 EARHWRPKSEVRGTQIGSYFGASLCSVDVDRDGSTDLVLIGAPHYYEQTRGGQVSVFPVP 499

Qy        482 RGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFH 541
              | | ||||:| |:|||| |||||| |||||||||| | ||||||||||::|||||:||
Db        500 -GVRGRWQCEATLHGEQGHPWGRFGVALTVLGDVNGDNLADVAIGAPGEEESRGAVYIFH 558

Qy        542 GTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQP 601
              | |   | || |||: ||:|| |||||||||||||||| |||||| ||||||||||| |
Db        559 GASRLEIMPSPSQRVTGSQLSLRLQYFGQSLSGGQDLTQDGLVDLAVGAQGHVLLLRSLP 618
```

```
Qy        602 VLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVTY 661
              :|:|:  : | | |||: |::| ::     ||||  ||| | | : | |  | :|  | |
Db        619 LLKVELSIRFAPMEVAKAVYQCWERTPTVLEAGEATVCLTVHKGSPDLL--GNVQGSVRY 676

Qy        662 DLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNFS 721
              ||||| ||  |||:|:|||| |    : |||   |||:|| ||:|:|| ||||:||||||
Db        677 DLALDPGRLISRAIFDETKNCTLTGRKTLGLGDHCETVKLLLPDCVEDAVSPIILRLNFS 736

Qy        722 LVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVGG 781
              ||    |  || |||  :|   ||  ||||||  : :|: || |:|:|  |  |||||
Db        737 LVRDSASP-RNLHPVLAVGSQDHITASLPFEKNCKQELLCEGDLGISFNFSGLQVLVVGG 795

Qy        782 PREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEVS 841
              |  |||||  |:||||| | | |::|  ||||:|:  | |  |   |||||:  : :
Db        796 SPELTVTVTVWNEGEDSYGTLVKFYYPAGLSYRRVTGTQ-QPHQYPLRLACEAEPAAQED 854

Qy        842 GALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEFQ 901
              |:|:|||||||| | :: || |||||  || ||::|||:|  :|||| | |||| ||
Db        855 --LRSSSCSINHPIFREGAKTTFMITFDVSYKAFLGDRLLLRAKASSENNKPDTNKTAFQ 912

Qy        902 LELPVKYAVYMVVTSHGVSTKYLNFTASENTSR-VMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||| || :::    || ::||::|    |    |:|:|:||    | : : | |||
Db        913 LELPVKYTVYTLISRQEDSTNHVNFSSSHGGRRQEAAHRYRVNNLSPLKLAVRVNFWVPV 972

Qy        961 RLNQTVIWDRPQVTFSENLSS-TCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              ||   :||  || |      :| :: : | : ||| ::::  |::|||| |   :|||
Db        973 LLNGVAVWD---VTLSSPAQGVSCVSQMKPPQNPDFLTQIQRRSVLDCSIADCLHFRCDI 1029

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              |   ||:| :  |:|||| | :|    :|:|| ||| |: ||:: |||| ||:|:| |
Db       1030 PSLDIQDELDFILRGNLSFGWVSQTLQEKVLLVSEAEITFDTSVYSQLPGQEAFLRAQVE 1089

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMM 1126
              | :| : |  |: |: |||||||||||||  ||||||||||||:|:
Db       1090 TTLEEYVVYEPIFLVAGSSVGGLLLLALITVVLYKLGFFKRQYKEML 1136
```

```
RESULT 11
ITAL_BOVIN
ID   ITAL_BOVIN     STANDARD;      PRT;  1165 AA.
AC   P61625;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE   alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE   chain) (CD11a).
GN   Name=ITGAL;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=14697514; DOI=10.1016/j.gene.2003.09.043;
RA   Fett T., Zecchinon L., Baise E., Desmecht D.;
RT   "The bovine (Bos taurus) CD11a-encoding cDNA: molecular cloning,
RT   characterisation and comparison with the human and murine
RT   glycoproteins.";
RL   Gene 325:97-101(2004).
CC   -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC       ICAM3 and ICAM4. It is involved in a variety of immune phenomena
CC       including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC       mediated killing, and antibody dependent killing by granulocytes
CC       and monocytes (By similarity).
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC       associates with beta-2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY267467; AAP94035.1; -.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     23       Potential.
FT   CHAIN        24   1165       Integrin alpha-L.
FT   DOMAIN       24   1084       Extracellular (Potential).
FT   TRANSMEM   1085   1105       Potential.
FT   DOMAIN     1106   1165       Cytoplasmic (Potential).
FT   REPEAT       40     88       FG-GAP 1.
FT   REPEAT       89    146       FG-GAP 2.
FT   DOMAIN      153    324       VWFA.
FT   REPEAT      347    397       FG-GAP 3.
FT   REPEAT      398    452       FG-GAP 4.
FT   REPEAT      454    513       FG-GAP 5.
FT   REPEAT      515    572       FG-GAP 6.
FT   REPEAT      575    627       FG-GAP 7.
FT   CA_BIND     465    473       Potential.
FT   CA_BIND     527    535       Potential.
FT   CA_BIND     587    595       Potential.
FT   SITE       1111   1115       GFFKR motif.
FT   DISULFID     71     78       By similarity.
FT   DISULFID    108    126       By similarity.
FT   DISULFID    650    704       By similarity.
FT   DISULFID    768    774       By similarity.
FT   DISULFID    842    858       By similarity.
FT   DISULFID    994   1009       By similarity.
FT   DISULFID   1017   1048       By similarity.
FT   CARBOHYD     33     33       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     86     86       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    185    185       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    646    646       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    667    667       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    723    723       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    859    859       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    894    894       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    929    929       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1056   1056       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1067   1067       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   1165 AA;  128725 MW;  DAEB3A3F1E1463CB CRC64;
```

```
  Query Match             26.9%;  Score 1583;  DB 1;  Length 1165;
  Best Local Similarity   35.1%;  Pred. No. 1.4e-95;
  Matches  415;  Conservative  195;  Mismatches  466;  Indels  106;  Gaps   33;

Qy          1 FNLDTENAMTFQENARG--FGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58
              :|||  :   |      |  ||  |:|: |: |||||| |    |  |:||||   || |
Db         24 YNLDVRHVQNFSFPLAGRHFGYRVLQV-GNGVVVGAPSE---GNSMGNLYQCQPETGDCL 79

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQP 118
              |: |       :  ||::||   :    |||| | : :|| :| |: |||:|   |||
Db         80 PVTLS--SNYTSKYLGMTLATDPTSDNLLACDPGLSRTCDQNIYLSGLCYLIHENLRGPV 137

Qy        119 QKFPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSE 178
              :   : | : : |: || ||| |:   :| :  :|:  ||::|  |    |: :|:|
Db        138 LQGHPGYQECIKGNVDLVFLFDGSMSLQQDEFEKIVDFMKDVMKKLSNSSYQFAAVQFST 197

Qy        179 EFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILI 238
              ||  ||| ::   :| :|:  :  :   |:|   |  | :|:|    ||| :| |:||
Db        198 YFRTEFTFLDYIRQKDPDALLAGVKHMRLLTNTFGAINYVAKEVFRPDLGARPDATKVLI 257
```

```
Qy        239 LITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQ 298
              :|||||         |  |  |  : :|||:||:|  |::::|:: |:  |||| : |
Db        258 IITDGEATD-----EHNIDAA--KDIIRYIIGIGKNFKTKESQEALHQFASKPVEEFVKI 310

Qy        299 INNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDW 358
              ::  ||  ||  :  :|::||: ||||      :||  |:|  | ||  ::     ::  ||: ||
Db        311 LDTFEKLKDLFTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSEGHGVVGAVGAKDW 370

Qy        359 AGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQHIGLV 416
              |||   |    : |||:    :  :    ||||    :  |  :  |  |||||||:| |
Db        371 AGGFLDLKADLKSSTFVGNEPLTVESRAGYLGYTVTWLPSRGTMSLLATGAPRYQHVGRV 430

Qy        417 AMFRQ--NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQ 474
               :|:|    | |     : | |||:|||  || |||| :| |:|:|| || || : |||:
Db        431 LLFQQPKRGGPWSQIQEIDGIQIGSYFGGELCGVDVDRDGETELLLIAAPLYYGEQRGGR 490

Qy        475 VSVCPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNR 534
              | :    :  :  :|  : | || | | ||||||:  | |:|||:|||||:||| ||  :
Db        491 VFIY---QKIQLEFQMVSELQGETGYPLGRFGAAIAALTDINGDELTDVAVGAPLEE--Q 545

Qy        535 GAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHV 594
              ||||:|:|  | |:||  |||| |:::   :|:||:|: | :||  ||| |: |||:| |
Db        546 GAVYIFNGQQG-GLSPRPSQRIEGTQMFSGIQWFGRSIHGVKDLGGDGLADVAVGAEGQV 604

Qy        595 LLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAG-EVRVCLHVQK--STRDRLR 651
              ::| |:||: :   : |:| |:  :  ||:       |: |  : ||  |:   ||
Db        605 IVLSSRPVVDIITSVSFSPAEIPVHEVECSYSTSNQKKEGVNLTVCFQVKSLIST----F 660

Qy        652 EGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPV 711
              :| : : :|| | ||   |  || :|   |:     | |  : ::|       | ||:| :
Db        661 QGHLVANLTYTLQLDGHRTRSRGLFPGGKHKLIGNTAVTPV-KSCFVFWFHFPICIQDLI 719

Qy        712 SPIVLRLNFSL---VGTPLS--AFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLS 766
              ||| : |::||    |||   |     | : : :  | |   |||||||  |  |: ||
Db        720 SPINVSLSYSLWEEEGTPRDPRALDRDIPPILKPSPHLETKEIPPEKNCGEDKNCEADLK 779

Qy        767 ITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQR 826
              : || |    | :      :| :|:||  ||:|  |||   ||  ||:||| |   :
Db        780 LAFSDMRSKILRLTPSASLSVRLTLRNTAEDAYWVQVTLSFPQGLSFRKVEIL---KPHS 836

Qy        827 SWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANV 886
              : ||       |  : :: ||::: ||| |:| |   : |:   | | |: : |:|||
Db        837 HVPVGCEELPEEAVVHS-RALSCNVSSPIFGEDSMVDIQVMFNTLQKGSWGDFIELQANV 895

Qy        887 TS--------ENNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQH 938
              :        |:|   |:       :|| | : ::      || |::||        ::|
Db        896 SCNNEDSSLLEDNSATTS-------IPVMYPINVLTKDQENSTLYISFTPKSPKIHHVKH 948

Qy        939 QYQV----SNLGQRSLPISLVFLVPVRLNQTVI---WD---RPQVTFS-ENLSSTCHTKE 987
               |||    ||      |:  :  ||   :: :|   |     | |  |  || |      |
Db        949 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPRNLESPSDEAE 1007

Qy        988 RLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHN 1047
                          :||       :| | |   ::|    : | :     || | :
Db       1008 -----------------SCSFGT--EFRCPIDF---RQEILVQVNGMVELRGTIKAS-S 1043

Qy       1048 HLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLAL 1107
               | : |:  | || |    | |: | : :|   ||:       | | | | :|||||| |
Db       1044 MLSLCSSLAISFNSSKHFHLYGRNASM-AQVVMKVDLVYEKEMLYLYVLSGIGGLLLLFL 1102

Qy       1108 ITAALYKLGFFKRQYKDMM-----------SEGGPPGAEPQ 1137
              |  ||||:|||||  |: |             : | |  | :
Db       1103 IFIALYKVGFFKRNLKEKMEANVDASSEIPGEDAGQPELEKE 1144
```

RESULT 12
ITAL_HUMAN
ID ITAL_HUMAN STANDARD; PRT; 1170 AA.
AC P20701; O43746;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE chain) (CD11a).
GN Name=ITGAL; Synonyms=CD11A;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RX MEDLINE=89139587; PubMed=2537322;
RA Larson R.S., Corbi A.L., Berman L., Springer T.;
RT "Primary structure of the leukocyte function-associated molecule-1
RT alpha subunit: an integrin with an embedded domain defining a protein
RT superfamily.";
RL J. Cell Biol. 108:703-712(1989).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=99425270; PubMed=10493829;
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from
RT human chromosome 16p and 16q.";
RL Genomics 60:295-308(1999).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.
RX MEDLINE=96036067; PubMed=7479767;
RA Qu A., Leahy D.J.;
RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L
RT beta 2) integrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.
RX MEDLINE=96398682; PubMed=8805579;
RA Qu A., Leahy D.J.;
RT "The role of the divalent cation in the structure of the I domain from
RT the CD11a/CD18 integrin.";
RL Structure 4:931-942(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.
RX MEDLINE=99425288; PubMed=10493852;
RA Kallen J., Welzenbach K., Ramage P., Geyl D., Kriwacki R., Legge G.,
RA Cottens S., Weitz-Schmidt G., Hommel U.;
RT "Structural basis for LFA-1 inhibition upon lovastatin binding to the
RT CD11a I-domain.";
RL J. Mol. Biol. 292:1-9(1999).
CC -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC ICAM3 and ICAM4. It is involved in a variety of immune phenomena
CC including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC mediated killing, and antibody dependent killing by granulocytes
CC and monocytes.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC associates with beta-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P20701-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P20701-2; Sequence=VSP_002738;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Leukocytes.
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11a.htm".
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Y00796; CAA68747.1; -.
DR   EMBL; AC002310; AAC31672.1; -.
DR   PIR; S03308; S03308.
DR   PDB; 1CQP; X-ray; A/B=153-334.
DR   PDB; 1DGQ; NMR; A=149-336.
DR   PDB; 1LFA; X-ray; A/B=-.
DR   PDB; 1MJN; X-ray; A=153-331.
DR   PDB; 1MQ8; X-ray; B/D=155-331.
DR   PDB; 1MQ9; X-ray; A=152-330.
DR   PDB; 1MQA; X-ray; A=152-330.
DR   PDB; 1ZON; X-ray; @=150-336.
DR   PDB; 1ZOO; X-ray; A/B=150-336.
DR   PDB; 1ZOP; X-ray; A/B=150-336.
DR   Genew; HGNC:6148; ITGAL.
DR   MIM; 153370; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   3D-structure; Alternative splicing; Calcium; Cell adhesion;
KW   Direct protein sequencing; Glycoprotein; Integrin; Magnesium;
KW   Receptor; Repeat; Signal; Transmembrane.
FT   SIGNAL       1     25
FT   CHAIN       26   1170       Integrin alpha-L.
FT   DOMAIN      26   1090       Extracellular (Potential).
FT   TRANSMEM  1091   1111       Potential.
FT   DOMAIN    1112   1170       Cytoplasmic (Potential).
FT   REPEAT      42     91       FG-GAP 1.
FT   REPEAT      92    149       FG-GAP 2.
FT   DOMAIN     156    327       VWFA.
FT   REPEAT     350    400       FG-GAP 3.
FT   REPEAT     401    455       FG-GAP 4.
FT   REPEAT     457    516       FG-GAP 5.
FT   REPEAT     518    575       FG-GAP 6.
FT   REPEAT     578    630       FG-GAP 7.
FT   CA_BIND    468    476       Potential.
FT   CA_BIND    530    538       Potential.
FT   CA_BIND    590    598       Potential.
FT   SITE      1115   1119       GFFKR motif.
FT   DISULFID    73     80       By similarity.
FT   DISULFID   111    129       By similarity.
FT   DISULFID   653    707       By similarity.
FT   DISULFID   771    777       By similarity.
FT   DISULFID   845    861       By similarity.
FT   DISULFID   998   1013       By similarity.
FT   DISULFID  1021   1052       By similarity.
FT   CARBOHYD    65     65       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   188    188       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   649    649       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   670    670       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   726    726       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   730    730       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   862    862       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   885    885       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   897    897       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1060   1060       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1071   1071       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC   954    954       Q -> QGVHGLVEMQTSKQILCRPAGDAEHTVGAQEGELPC
FT                               PWGVSEAFRDNIRAGPCR (in isoform 2).
FT                               /FTId=VSP_002738.
FT   CONFLICT   214    214       R -> W (in Ref. 1 and 2).
FT   CONFLICT   660    660       Y -> I (in Ref. 2).
FT   STRAND     155    162
FT   STRAND     164    164
FT   TURN       165    166
FT   HELIX      169    185
FT   TURN       186    188
FT   STRAND     191    198
FT   STRAND     202    206
FT   HELIX      208    214
FT   HELIX      217    221
FT   TURN       222    223
FT   STRAND     229    229
FT   HELIX      233    243
FT   TURN       244    244
FT   HELIX      247    249
FT   TURN       250    250
FT   TURN       253    254
FT   STRAND     256    263
FT   HELIX      274    276
FT   TURN       277    278
FT   STRAND     280    286
FT   HELIX      288    290
FT   HELIX      293    297
FT   TURN       298    299
FT   HELIX      300    302
FT   HELIX      307    310
FT   STRAND     311    314
FT   HELIX      317    319
FT   HELIX      323    330
SQ   SEQUENCE  1170 AA;  128819 MW;  39A7AF92EF286FC0 CRC64;

  Query Match             26.5%;  Score 1557.5;  DB 1;  Length 1170;
  Best Local Similarity   34.7%;  Pred. No. 7.1e-94;
  Matches  407;  Conservative  206;  Mismatches  460;  Indels  101;  Gaps   37;

Qy          1 FNLDTENAMTFQ--ENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58
              :|||   | :|       | ||  |:|: |: |:|||| |    |  ||||||  || |
Db         26 YNLDVRGARSFSPPRAGRHFGYRVLQV-GNGVIVGAPGE---GNSTGSLYQCQSGTGHCL 81

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLR--- 115
              |: |:      :  ||::||   :   :||| | : :|| :|||: |||:|| ||:
Db         82 PVTLR-GSNYTSKYLGMTLATDPTDGSILACDPGLSRTCDQNTYLSGLCYLFRQNLQGPM 140

Qy        116 -QQPQKFPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLM 174
               |    | | ::|    : |: || ||| |: | :|::  :|:  ||::|  :  |: :
Db        141 LQGRPGFQECIKG----NVDLVFLFDGSMSLQPDEFQKILDFMKDVMKKLSNTSYQFAAV 196

Qy        175 QYSEEFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAF 234
              |:|  ::  | | ::    :| :|:| :  :|  |:|   |  |  |:|    ||| :|
Db        197 QFSTSYKTEFDFSDYVKRKDPDALLKHVKHMLLLTNTFGAINYVATEVFREELGARPDAT 256

Qy        235 KILILITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRD 294
              |:||:|||||       | |   :     :|||:||:|  |::::|:: |:  |||| :
Db        257 KVLIIITDGE--ATDSGNIDAAKD-----IIRYIIGIGKHFQTKESQETLHKFASKPASE 309

Qy        295 HVFQINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVG 354
              |  :: || || :  :|::||: ||||     :||  |:|  | || ::    ::  ||
Db        310 FVKILDTFEKLKDLFTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSRGHAVVGAVG 369

Qy        355 SYDWAGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQH 412
              : |||||   |    :  |||    :  ::   ||||    :  | :   |  |||||||
Db        370 AKDWAGGFLDLKADLQDDTFIGNEPLTPEVRAGYLGYTVTWLPSRQKTSLLASGAPRYQH 429

Qy        413 IGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQT 470
              :| | :|:  |  | |      : |||||:|||  || |||| :| |:|:|||||  :| :
Db        430 MGRVLLFQEPQGGGHWSQVQTIHGTQIGSYFGGELCGVDVDQDGETELLLIGAPLFYGEQ 489

Qy        471 RGGQVSVCPLPRGQRARWQCDAV--LYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAP 528
              |||:| :      || :   : |  | |: | | |||| |:| | |:||| | |||:|||
```

```
Db        490 RGGRVFIY-----QRRQLGFEEVSELQGDPGYPLGRFGEAITALTDINGDGLVDVAVGAP 544

Qy        529 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTV 588

Db        545 LEE--QGAVYIFNGRHG-GLSPQPSQRIEGTQVLSGIQWFGRSIHGVKDLEGDGLADVAV 601

Qy        589 GAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVV-KGKEAGEVRVCLHVQKSTR 647

Db        602 GAESQMIVLSSRPVVDMVTLMSFSPAEIPVHEVECSYSTSNKMKEGVNITICFQI-KSLY 660

Qy        648 DRLREGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCI 707

Db        661 PQF-QGRLVANLTYTLQLDGHRTRRRGLFPGGRHELRRNIAVT-TSMSCTDFSFHFPVCV 718

Qy        708 EDPVSPIVLRLNFSL---VGTPLS--AFGN-----LRPVLAEDAQRLFTALFPFEKNCGN 757

Db        719 QDLISPINVSLNFSLWEEEGTPRDQRAQGKDIPPILRPSLHSETWEI-----PFEKNCGE 773

Qy        758 DNICQDDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVS 817

Db        774 DKKCEANLRVSFSPARSRALRLTAFASLSVELSLSNLEEDAYWVQLDLHFPPGLSFRKVE 833

Qy        818 TLQNQRSQRSWRLACES--ASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKAS 875

Db        834 ML---KPHSQIPVSCEELPEESRLLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNSS 887

Qy        876 LGNKLLLKANVTSENN----MPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASEN 931

Db        888 WGDSVELHANVTCNNEDSDLLEDNSATTI---IPILYPINILIQDQEDSTLYVSFTPKGP 944

Qy        932 TSRVMQHQYQV---SNLGQRSLP-ISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTK- 986

Db        945 KIHQVKHMYQVRIQPSIHDHNIPTLEAVVGVPQPPSEGPITHQWSVQMEPPV--PCHYED 1002

Qy        987 -ERLPSHSD--FLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIK 1043

Db       1003 LERLPDAAEPCLPGALFRCPVV-----------------FRQEILVQVIGTLELVGEIE 1044

Qy       1044 TSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLL 1103

Db       1045 AS-SMFSLCSSLSISFNSSKHFHLYGSNASL-AQVVMKVDVVYEKQMLYLYVLSGIGGLL 1102

Qy       1104 LLALITAALYKLGFFKRQYKDMMSEG-GPPGAEP 1136

Db       1103 LLLLIFIVLYKVGFFKRNLKEKMEAGRGVPNGIP 1136
```

```
RESULT 13
Q6TYB8
ID   Q6TYB8      PRELIMINARY;      PRT;  1166 AA.
AC   Q6TYB8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR   EMBL; AY382558; AAQ90015.1; -.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 1.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Cell adhesion; Integrin; Transmembrane.
SQ   SEQUENCE   1166 AA;  128723 MW;  7B8D8AFBA896C9DF CRC64;
```

```
  Query Match           26.5%;  Score 1557;  DB 2;  Length 1166;
  Best Local Similarity 34.6%;  Pred. No. 7.6e-94;
  Matches  409;  Conservative  198;  Mismatches  469;  Indels  106;  Gaps   33;

Qy          1 FNLDTENAMTFQENARG--FGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58

Db         25 YNLDVRHVQNFSFPLAGRHFGYRVLQV-GNGVVVGAPSE---GNSMGNLYQCQPETGDCL 80

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQP 118

Db         81 PVTLS--SNYTSKYLGMTLATDPTSDNLLACDPGLSRTCDQNIYLSGLCYLIHENLRGPV 138

Qy        119 QKFPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSE 178

Db        139 LQGHPGYQECIKGNVDLVFLFDGSMSLQQDEFEKIVDFMKDVMKKLSNSSYQFAAVQFST 198

Qy        179 EFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILI 238

Db        199 YFRTEFTFLDYIKQKDPDALLAGVKHMRLLTNTFGAINYVAKEVFRPDLGARPDATKVLI 258

Qy        239 LITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQ 298

Db        259 IITDGK---PPTNTTLMRPKTSS----RSLLGIGKNFKTKESQEALHQFASKPVEEFVKI 311

Qy        299 INNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDW 358

Db        312 LDTFEKLKDLFTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSEGHGVVGAVGAKDW 371

Qy        359 AGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQHIGLV 416

Db        372 AGGFLDLKADLKSSTFVGNEPLTVESRAGYLGYTVTRLPSRGTMSLLATGAPKYQHVGRV 431

Qy        417 AMFRQ--NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQ 474

Db        432 LLFQQPKRGGPWSQIQEIDGIQIGSYFGGELCGVDVDRDGETELLLIAAPLYYGEQRGGR 491

Qy        475 VSVCPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNR 534

Db        492 VFIY---QKIQLEFQMVSELQGETGYPLGRFGAAIAALTDINGDELTDVAVGAPLEE--Q 546

Qy        535 GAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHV 594

Db        547 GAVYIFNGQQG-GLSPRPSQRIEGTQMFSGIQWFGRSIHGVKDLGGDGLADVAVGAEGQV 605

Qy        595 LLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAG-EVRVCLHVQK--STRDRLR 651

Db        606 IVLSSRPVVDIITSVSFSPAEIPVHEVECSYSTSNQKKEGVNLTVCFQVKSLIST----F 661

Qy        652 EGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPV 711

Db        662 QGHLVANLTYTLQLDGHRTRSRGLFPGGKHKLIGNTAVTPV-KSCFVFWFHFPICIQDLI 720

Qy        712 SPIVLRLNFSL---VGTPLS--AFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLS 766

Db        721 SPINVSLSYSLWEEEGTPRDPRALDRDIPPILKPSPHLETKEIPFEKNCGEDKNCEADLK 780

Qy        767 ITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQR 826

Db        781 LAFSDMRSKILRLTPSASLSVRLTLRNTAEDAYWVQVTLSFPQGLSFRKVEIL---KPHS 837

Qy        827 SWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANV 886

Db        838 HVPVGCEELPEEAVVHS-RALSCNVSSPIFGEDSMVDIQVMFNTLQKGSWGDFIELQANV 896

Qy        887 TS--------ENNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQH 938
```

```
Db        897 SCNNEDSSLLEDNSATTS-------IPVMYPINVLTKDQENSTLYISFTPKSPRIHHVKH 949

Qy        939 QYQV----SNLGQRSLPISLVFLVPVRLNQTVI---WD---RPQVTFS-ENLSSTCHTKE 987

Db        950 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPRNLESPSDEAE 1008

Qy        988 RLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHN 1047

Db       1009 ------------------SCSFGT--EFRCPIDF---RQEILVQVNGMVELRGTIKAS-S 1044

Qy       1048 HLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLAL 1107

Db       1045 MLSLCSSLAISFNSSKHFHLHGRNASM-AQVVMKVDLVYEKEMLYLYVLSGIGGLLLLFL 1103

Qy       1108 ITAALYKLGFFKRQYKDMM------------SEGGPPGAEPQ 1137

Db       1104 IFIALYKVGFFKRNLKEKMEANVDASSEIPGEDAGQPELEKE 1145


RESULT 14
AAQ90015
ID   AAQ90015    PRELIMINARY;      PRT;  1166 AA.
AC   AAQ90015;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT   "Molecular cloning and sequencing of bovine CD11a.";
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY382558; AAQ90015.1; -.
SQ   SEQUENCE   1166 AA;  128723 MW;  7B8D8AFBA896C9DF CRC64;

  Query Match             26.5%;  Score 1557;  DB 2;  Length 1166;
  Best Local Similarity   34.6%;  Pred. No. 7.6e-94;
  Matches  409;  Conservative  198;  Mismatches  469;  Indels  106;  Gaps   33;

Qy          1 FNLDTENAMTFQENARG--FGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58

Db         25 YNLDVRHVQNFSFPLAGRHFGYRVLQV-GNGVVVGAPSE---GNSMGNLYQCQPETGDCL 80

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQP 118

Db         81 PVTLS--SNYTSKYLGMTLATDPTSDNLLACDPGLSRTCDQNIYLSGLCYLIHENLRGPV 138

Qy        119 QKFPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSE 178

Db        139 LQGHPGYQECIKGNVDLVFLFDGSMSLQQDEFEKIVDFMKDVMKKLSNSSYQFAAVQFST 198

Qy        179 EFRIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILI 238

Db        199 YFRTEFTFLDYIKQKDPDALLAGVKHMRLLTNTFGAINYVAKEVFRPDLGARPDATKVLI 258

Qy        239 LITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQ 298

Db        259 IITDGK---PPTNTTLMRPKTSS----RSLLGIGKNFKTKESQEALHQFASKPVEEFVKI 311

Qy        299 INNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDW 358

Db        312 LDTFEKLKDLFTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSEGHGVVGAVGAKDW 371

Qy        359 AGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQHIGLV 416

Db        372 AGGFLDLKADLKSSTFVGNEPLTVESRAGYLGYTVTRLPSRGTMSLLATGAPKYQHVGRV 431

Qy        417 AMFRQ--NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQ 474

Db        432 LLFQQPKRGGPWSQIQEIDGIQIGSYFGGELCGVDVDRDGETELLLIAAPLYYGEQRGGR 491

Qy        475 VSVCPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNR 534

Db        492 VFIY---QKIQLEFQMVSELQGETGYPLGRFGAAIAALTDINGDELTDVAVGAPLEE--Q 546

Qy        535 GAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHV 594

Db        547 GAVYIFNGQQG-GLSPRPSQRIEGTQMFSGIQWFGRSIHGVKDLGGDGLADVAVGAEGQV 605

Qy        595 LLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAG-EVRVCLHVQK--STRDRLR 651

Db        606 IVLSSRPVVDIITSVSFSPAEIPVHEVECSYSTSNQKKEGVNLTVCFQVKSLIST----F 661

Qy        652 EGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPV 711

Db        662 QGHLVANLTYTLQLDGHRTRSRGLFPGGKHKLIGNTAVTPV-KSCFVFWFHFPICIQDLI 720

Qy        712 SPIVLRLNFSL---VGTPLS--AFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLS 766

Db        721 SPINVSLSYSLWEEEGTPRDPRALDRDIPPILKPSPHLETKEIPFEKNCGEDKNCEADLK 780

Qy        767 ITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQR 826

Db        781 LAFSDMRSKILRLTPSASLSVRLTLRNTAEDAYWVQVTLSFPQGLSFRKVEIL---KPHS 837

Qy        827 SWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANV 886

Db        838 HVPVGCEELPEEAVVHS-RALSCNVSSPIFGEDSMVDIQVMFNTLQKGSWGDFIELQANV 896

Qy        887 TS--------ENNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQH 938

Db        897 SCNNEDSSLLEDNSATTS-------IPVMYPINVLTKDQENSTLYISFTPKSPRIHHVKH 949

Qy        939 QYQV----SNLGQRSLPISLVFLVPVRLNQTVI---WD---RPQVTFS-ENLSSTCHTKE 987

Db        950 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPRNLESPSDEAE 1008

Qy        988 RLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHN 1047

Db       1009 ------------------SCSFGT--EFRCPIDF---RQEILVQVNGMVELRGTIKAS-S 1044

Qy       1048 HLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLAL 1107

Db       1045 MLSLCSSLAISFNSSKHFHLHGRNASM-AQVVMKVDLVYEKEMLYLYVLSGIGGLLLLFL 1103

Qy       1108 ITAALYKLGFFKRQYKDMM------------SEGGPPGAEPQ 1137

Db       1104 IFIALYKVGFFKRNLKEKMEANVDASSEIPGEDAGQPELEKE 1145


RESULT 15
ITAL_MOUSE
ID   ITAL_MOUSE      STANDARD;      PRT;  1163 AA.
AC   P24063;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE   alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE   chain) (CD11a).
GN   Name=Itgal; Synonyms=Lfa-1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91268576; PubMed=2051027;
RA   Kaufmann Y., Tseng E., Springer T.A.;
RT   "Cloning of the murine lymphocyte function-associated molecule-1
```

```
RT   alpha-subunit and its expression in COS cells.";
RL   J. Immunol. 147:369-374(1991).
RN   [2]
RP   SEQUENCE OF 24-42.
RX   MEDLINE=85188276; PubMed=3887182;
RA   Springer T.A., Teplow D.B., Dreyer W.J.;
RT   "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
RT   glycoproteins and unexpected relation to leukocyte interferon.";
RL   Nature 314:540-542(1985).
CC   -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC       ICAM3 and ICAM4. Is involved in a variety of immune phenomena
CC       including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC       mediated killing, and antibody dependent killing by granulocytes
CC       and monocytes. Mice expressing a null mutation of the alpha-L
CC       subunit gene demonstrate impaired tumor rejection and impaired
CC       leukocytes recruitment.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Leukocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M60778; AAA39426.1; -.
DR   PIR; I56126; I56126.
DR   HSSP; P20701; 1DGQ.
DR   MGD; MGI:96606; Itgal.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 2.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Direct protein sequencing; Glycoprotein;
KW   Integrin; Magnesium; Receptor; Repeat; Signal; Transmembrane.
FT   SIGNAL        1     23
FT   CHAIN        24   1163       Integrin alpha-L.
FT   DOMAIN       24   1087       Extracellular (Potential).
FT   TRANSMEM   1088   1108       Potential.
FT   DOMAIN     1109   1163       Cytoplasmic (Potential).
FT   REPEAT       39     88       FG-GAP 1.
FT   REPEAT       89    146       FG-GAP 2.
FT   DOMAIN      153    325       VWFA.
FT   REPEAT      348    398       FG-GAP 3.
FT   REPEAT      399    454       FG-GAP 4.
FT   REPEAT      455    514       FG-GAP 5.
FT   REPEAT      516    573       FG-GAP 6.
FT   REPEAT      576    628       FG-GAP 7.
FT   CA_BIND     466    474       Potential.
FT   CA_BIND     528    536       Potential.
FT   CA_BIND     588    596       Potential.
FT   SITE       1111   1115       GFFKR motif.
FT   DISULFID     70     77       By similarity.
FT   DISULFID    108    126       By similarity.
FT   DISULFID    147    199       By similarity.
FT   DISULFID    651    705       By similarity.
FT   DISULFID    767    773       By similarity.
FT   DISULFID    840    856       By similarity.
FT   DISULFID    993   1009       By similarity.
FT   DISULFID   1017   1048       By similarity.
FT   CARBOHYD     86     86       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    185    185       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    270    270       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    444    444       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    668    668       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    696    696       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    724    724       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    728    728       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    776    776       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    857    857       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    880    880       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    890    890       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    899    899       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    927    927       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1056   1056       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   1163 AA;  128343 MW;  A7A3078489E8232F CRC64;
```

Query Match 26.2%; Score 1538.5; DB 1; Length 1163;
Best Local Similarity 34.2%; Pred. No. 1.3e-92;
Matches 401; Conservative 214; Mismatches 458; Indels 99; Gaps 36;

```
Qy          1 FNLDTENAMTFQENA-RGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEP 59
              :||||    :|  | | ||  |:|::   ||||||  |    :  | || |  |:  |:|
Db         24 YNLDTRPTQSFLAQAGRHFGYQVLQIEDG-VVVGAPGE---GDNTGGLYHCRTSSEFCQP 79

Qy         60 IRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQ 119
              : |       :  ||::||   :   ||||  | : :|| :|||: |||:||  :|
Db         80 VSLH-GSNHTSKYLGMTLATDAAKGSLLACDPGLSRTCDQNTYLSGLCYLFPQSLEGPML 138

Qy        120 KFPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEE 179
              :  | : | :   |: || ||| |:   || :  ||:  || :|  :   |: :|:| :
Db        139 QNRPAYQECMKGKVDLVFLFDGSQSLDRKDFEKILEFMKDVMRKLSNTSYQFAAVQFSTD 198

Qy        180 FRIHFTFKEF-QNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILI 238
              |  ||| :: : | ||  |:  :  :   |:|   |  ||  :|   :||| :| |:|:
Db        199 CRTEFTFLDYVKQNKNPDVLLGSVQPMFLLTNTFRAINYVVAHVFKEESGARPDATKVLV 258

Qy        239 LITDGEKFGDPLGYEDVIPEADREG-------VIRYVIGVGDAFRSEKSRQELNTVASKP 291
              :||||              ||  :|       : ||:||:|  | | : :: |:  ||:|
Db        259 IITDG-------------EASDKGNISAAHDITRYIIGIGKHFVSVQKQKTLHIFASEP 304

Qy        292 PRDHVFQINNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLS 351
              :  |  :: || || :   |: :|:|||||      :||  |:|  | || ::    ::
Db        305 VEEFVKILDTFEKLKDLFTDLQRRIYAIEGTNRQDLTSFNMELSSSGISADLSKGHAVVG 364

Qy        352 TVGSYDWAGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPR 409
              ||: |||||  |  : :||:  : ||:  ||||  | :  |:   |  ||||
Db        365 AVGAKDWAGGFDLDLREDLQGATFVGQEPLTSDVRGGYLGYTVAWMTSRSSRPLLAAGAPR 424

Qy        410 YQHIGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYY 467
              |||:| | :|:  :  | |     ::|||||:|||  |||||:| :|  :|:|||||  ::
Db        425 YQHVGQVLLFQAPEAGGRWNQTQKIEGTQIGSYFGGELCSVDLDQDGEAELLLIGAPLFF 484

Qy        468 EQTRGGQVSVCPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGA 527
              : |||:|      : ::: ::  : | |: | | ||||||:| | \:|||:|||||:||
Db        485 GEQRGGRVFTY---QRRQSLFEMVSELQGDPGYPLGRFGAAITALTDINGDRLTDVAVGA 541

Qy        528 PGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLT 587
              | ||  :||||:|:|  | |:||  |||| |::: | :::||:|: | :||  | | |:
Db        542 PLEE--QGAVYIFNGKPG-GLSPQPSQRIQGAQVFPGIRWFGRSIHGVKDLGGDRLADVV 598

Qy        588 VGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAG-EVRVCLHVQKST 646
              |||:| |::| |:||:  |   : |:| |:  :  ||:    : :: | ::: |  ::  |
Db        599 VGAEGRVVVLSSRPVVDVVTELSFSPEEIPVHEVECSYSAREEQKHGVKLKACFRIKPLT 658

Qy        647 RDRLREGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNC 706
                :|:: : ::| | ||  |  || :| :  :      | :    ::|      | |
Db        659 PQ--FQGRLLANLSYTLQLDGHRMRSRGLFPDGSHELSGNTSITP-DKSCLDFHFHFPIC 715

Qy        707 IEDPVSPIVLRLNFSLV---GTPLSAFGN-LRPVLAEDAQRLFTALFPFEKNCGNDNICQ 762
```

```
Db        716 IQDLISPINVSLNFSLLEEEGTPRDQKGRAMQPILRPSIHTV-TKEIPFEKNCGEDKKCE 774

Qy        763 DDLSITFSFMSLDCLVVGGP------REFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKV 816

Db        775 ANLTLSSPARS-------GPLRLMSSASLAVEWTLSNSGEDAYWVRLDLDFPRGLSFRKV 827

Qy        817 STLQNQRSQRSWRLACESASSTEVSGAL-KSTSCSINHPIFPENSEVTFNITFDVDSKAS 875

Db        828 EMLQ---PHSRMPVSCEEL--TEGSSLLTKTLKCNVSSPIFKAGQEVSLQVMFNTLLNSS 882

Qy        876 LGNKLLLKANVTSEN-NMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSR 934

Db        883 WEDFVELNGTVHCENENSSLQEDNSAATHIPVLYPVNILTKEQENSTLYISFTPKGPKTQ 942

Qy        935 VMQHQYQVSNLGQRSLPISLVFLVPVRLNQTVIWDRPQ----VTFSENLSS----TCHTK 986

Db        943 QVQHVYQV-----RIQPSAYDHNMPT-LEALVGVPRPHSEDLITYTWSVQTDPLVTCHSE 996

Qy        987 E-RLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTS 1045

Db        997 DLKRPSSE---AEQPCLPGV--------QFRCPIVF---RWEILIQVTGTVELSKEIKAS 1042

Qy       1046 HNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLL 1105

Db       1043 -STLSLCSSLSVSFNSSKHFHLYGSKA-SEAQVLVKVDLIHEKEMLHVYVLSGIGGLVLL 1100

Qy       1106 ALITAALYKLGFFKRQYKDMM-SEGGPPGAEP 1136

Db       1101 FLIFLALYKVGFFKRNLKEKMEADGGVPNGSP 1132
```

Search completed: November 9, 2004, 12:17:38
Job time : 141.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:45:42 ; Search time 103.25 Seconds
(without alignments)
3950.365 Million cell updates/sec

Title: US-09-902-481B-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5862	99.8	1153	2	AAW65090	Aaw65090 Human Bet
2	5862	99.8	1153	3	AAB07360	Aab07360 Human CD1
3	5862	99.8	1153	5	AAU80252	Aau80252 Human int
4	5862	99.8	1153	5	ABG61469	Abg61469 Human Bet
5	5862	99.8	1153	5	AAO14428	Aao14428 Integrin
6	5862	99.8	1153	7	ADD25615	Add25615 Binding d
7	5852	99.6	1153	2	AAR04136	Aar04136 Alpha sub
8	5846.5	99.5	1152	8	ADM99589	Adm99589 Human int
9	5839.5	99.4	1152	8	ADP12435	Adp12435 Protein e
10	3483	59.3	1163	8	ADP44061	Adp44061 Human CD1
11	3469	59.0	1163	8	ADN02004	Adn02004 Human inf
12	3469	59.0	1163	8	ADQ17510	Adq17510 Human sof
13	3460	58.9	1163	2	AAR07120	Aar07120 p150.95 a
14	3446	58.6	1163	2	AAW65091	Aaw65091 Human Bet
15	3446	58.6	1163	3	AAB07361	Aab07361 Human CD1
16	3446	58.6	1163	5	ABG61470	Abg61470 Human Bet
17	3444	58.6	1163	6	ABU07406	Abu07406 Protein d
18	3444	58.6	1163	7	ADG32005	Adg32005 Human hom
19	3411	58.0	1161	2	AAR78166	Aar78166 Human bet
20	3411	58.0	1161	2	AAW23049	Aaw23049 Human bet
21	3411	58.0	1161	2	AAW57491	Aaw57491 Human bet
22	3411	58.0	1161	2	AAW65089	Aaw65089 Human Bet
23	3411	58.0	1161	2	AAW72825	Aaw72825 Human alp
24	3411	58.0	1161	2	AAW73342	Aaw73342 Human alp
25	3411	58.0	1161	3	AAB07359	Aab07359 Human alp
26	3411	58.0	1161	5	ABG61468	Abg61468 Human Bet
27	3395.5	57.8	1161	2	AAW23064	Aaw23064 Human bet
28	3395.5	57.8	1161	2	AAW65106	Aaw65106 Human Bet
29	3395.5	57.8	1161	2	AAW72837	Aaw72837 Human alp
30	3395.5	57.8	1161	2	AAW73343	Aaw73343 Human alp
31	3395.5	57.8	1161	3	AAB07376	Aab07376 Human alp
32	3395.5	57.8	1161	5	ABG61485	Abg61485 Human Bet
33	3234.5	55.0	1161	2	AAR78169	Aar78169 Rat alpha
34	3232.5	55.0	1161	2	AAW23062	Aaw23062 Rat beta
35	3232.5	55.0	1161	2	AAW60004	Aaw60004 Rat alpha
36	3232.5	55.0	1161	2	AAW72824	Aaw72824 Rat alpha
37	3232.5	55.0	1161	3	AAB07374	Aab07374 Rat alpha
38	3232.5	55.0	1161	5	ABG61483	Abg61483 Rat Beta2
39	3225.5	54.9	1161	2	AAW65104	Aaw65104 Rat beta-
40	3225.5	54.9	1161	2	AAW73345	Aaw73345 Rat alpha
41	3222	54.8	1161	2	AAW23061	Aaw23061 Mouse bet
42	3222	54.8	1161	2	AAW60003	Aaw60003 Mouse alp
43	3222	54.8	1161	2	AAW65103	Aaw65103 Mouse bet
44	3222	54.8	1161	2	AAW72836	Aaw72836 Mouse alp
45	3222	54.8	1161	2	AAW73347	Aaw73347 Mouse alp

ALIGNMENTS

RESULT 1
AAW65090
ID AAW65090 standard; protein; 1153 AA.
XX
AC AAW65090;
XX
DT 28-SEP-1998 (first entry)
XX
DE Human Beta-integrin CD11b subunit protein.
XX
KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN US5728533-A.
XX
PD 17-MAR-1998.
XX
PF 07-JUN-1995; 95US-00485618.
XX
PR 23-DEC-1993; 93US-00173497.
PR 05-AUG-1994; 94US-00286889.
PR 21-DEC-1994; 94US-00362652.
XX
PA (ICOS-) ICOS CORP.
XX
PI Van Der Vieren M, Gallatin WM;
XX
DR WPI; 1998-206565/18.
XX
PT Screening assay for modulators of integrin binding - using immobilised or
PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.
XX
PS Example 5; Fig 1A-D; 106pp; English.
XX
CC This sequence represents a human beta-integrin CD11b subunit which is
CC used to describe a method for identifying compounds that modulate the
CC interaction of the beta-integrin alpha-d subunit with a binding partner
CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha
CC -d binding partner, one of which is immobilised and the other of which is
CC labelled, in the presence of a test compound, and determining if the
CC compound affects binding between the alpha-d polypeptide and alpha-d
CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment
CC comprising the cytoplasmic, transmembrane or extracellular domain of
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

```
CC   diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC   asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC   and rheumatoid arthritis
XX
SQ   Sequence 1153 AA;

  Query Match             99.8%;  Score 5862;  DB 2;  Length 1153;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||| |||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 2
AAB07360
ID   AAB07360 standard; protein; 1153 AA.
XX
AC   AAB07360;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Human CD11b protein sequence.
XX
KW   Human; macrophage infiltration inhibition; alpha_d integrin;
KW   leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW   inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW   atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW   lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW   rheumatoid arthritis; central nervous system injury; CD11b.
XX
OS   Homo sapiens.
XX
PN   WO200029446-A1.
XX
PD   25-MAY-2000.
XX
PF   16-NOV-1999;   99WO-US027139.
XX
PR   16-NOV-1998;   98US-00193043.
PR   08-JUL-1999;   99US-00350259.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin MW,  Van Der Vieren M;
XX
DR   WPI; 2000-387751/33.
XX
PT   Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT   macrophage infiltration and reduce inflammation at central nervous system
PT   injury sites.
XX
PS   Example 5; Fig 1; 270pp; English.
XX
CC   Integrins are a class of membrane-associated molecules that participate
CC   in cellular adhesion. Integrins are made up of an alpha subunit and a
CC   beta subunit. One class of human integrins are restricted to expression
CC   in white blood cells and have a common beta2 subunit: the leukocyte
CC   integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC   have an important role in immune and inflammatory responses. The present
CC   protein sequence is the human integrin alpha subunit CD11b. This sequence
CC   was used in an alignment to identify a novel beta2 integrin alpha
CC   subunit: alpha_d (AAA60014 and AAB07359). The present sequence has
CC   approximately 60% identity to the protein sequence of alpha_d. The
CC   Alpha_d gene and protein may be useful in therapy for diseases linked to
CC   aberrant alpha_d function e.g. Type I diabetes, atherosclerosis, multiple
CC   sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
CC   distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
CC   (LAD). In addition, anti-alpha_d monoclonal antibodies may be used in the
CC   inhibition of macrophage infiltration at the site of a central nervous
CC   system injury. The monoclonal antibodies can also be used to detect and
CC   diagnose Crohn's disease
XX
```

SQ Sequence 1153 AA;

Query Match 99.8%; Score 5862; DB 3; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||||| |||:|||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||:|||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEPNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEPNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 3
AAU80252
ID AAU80252 standard; protein; 1153 AA.
XX
AC AAU80252;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human integrin 1 alpha-M subunit protein.
XX
KW Integrin; antiinflammatory; imunosuppression; nephritis; dermatitis;
KW inflammatory disease; autoimmune disorder; Crohn's disease;
KW human immunodeficiency virus; HIV; myocardial infarction;
KW Sjorgen's syndrome; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 499. .500
FT /note= "Encoded by GGG CAG AGG"
XX
PN WO200218583-A2.
XX
PD 07-MAR-2002.
XX
PF 31-AUG-2001; 2001WO-US027227.
XX
PR 01-SEP-2000; 2000US-0229700P.
XX
PA (BLOO-) CENT BLOOD RES INC.
XX
PI Springer TA, Shimoaka M, Lu C;
XX
DR WPI; 2002-382964/41.
DR N-PSDB; ABK50046.
XX
PT Modified integrin-I or integrin I-like domain polypeptide useful as an
PT immunogen to produce antibodies specific to polypeptide, comprises a
PT disulfide bond such that polypeptide is stabilized in a desired
PT conformation.
XX
PS Disclosure; Page 109-112; 112pp; English.
XX
CC This invention relates to a modified integrin-I or integrin I-like domain
CC polypeptide comprising at least one disulfide bond so that the domain is
CC stabilised in a desired conformation. The polypeptide of the invention
CC may have antiinflammatory or imunosuppressive activities. The
CC polypeptides of the invention have an open conformation and are useful as
CC immunogens to produce antibodies that selectively bind to integrin I-
CC domain; and for identifying a modulator of integrin activity, or of
CC interaction of an integrin and a cognate ligand. The polypeptide of the
CC invention, or antibodies (preferably anti-LFA-1 antibody) is useful for
CC treating or preventing an integrin mediated disorder which is an
CC inflammatory or autoimmune disorder in a subject and for inhibiting the
CC binding of an integrin to a cognate ligand such as Crohn's disease,
CC nephritis; human immunodeficiency virus (HIV), myocardial infarction,
CC Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic
CC composition comprising the peptide of the invention is useful for
CC treating an integrin mediated disorder in a subject. The polypeptides
CC and/or active or antigenic fragments are useful as reagents for diagnosis
CC of integrin-mediated disorders. The present sequence represents the human
CC integrin-1 alpha-M protein subunit used to generate the mutant
CC polypeptides of the invention

```
XX
SQ   Sequence 1153 AA;

  Query Match           99.8%;  Score 5862;  DB 5;  Length 1153;
  Best Local Similarity  99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||||| |||:|||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||::|||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 4
ABG61469
ID   ABG61469 standard; protein; 1153 AA.
XX
AC   ABG61469;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human Beta2 integrin alphaCD11b subunit.
XX
KW   Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; LAD;
KW   leukocyte adhesion deficiency; inflammatory response; diabetes;
KW   multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW   inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW   immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
KW   intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW   locomotor recovery; locomotor damage; locomotor impairment;
KW   autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX
OS   Homo sapiens.
XX
PN   WO200230980-A2.
XX
PD   18-APR-2002.
XX
PF   15-OCT-2001; 2001WO-US032059.
XX
PR   13-OCT-2000; 2000US-00688307.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin WM,  Van Der Vieren M;
XX
DR   WPI; 2002-463260/49.
XX
PT   Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT   recovery, inhibiting locomotor damage, limiting locomotor impairment, or
PT   limiting autonomic and sensory dysfunction following spinal cord injury.
XX
PS   Example 5; Page 191-194; 270pp; English.
XX
CC   The invention relates to promoting locomotor recovery, inhibiting
CC   locomotor damage, limiting locomotor impairment, or limiting autonomic
CC   and sensory dysfunction following spinal cord injury by administering an
CC   anti-alpha_d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
CC   spinal cord injury victim. The method also involves the use of a ligand
CC   selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,
CC   vascular cell adhesion molecule). The method is useful for promoting
CC   locomotor recovery, inhibiting locomotor damage, limiting locomotor
CC   impairment, or limiting autonomic and sensory dysfunction following
CC   spinal cord injury. In particular, the spinal cord injury comprises
CC   compression of the spinal cord. The antibodies are also useful for
CC   reducing inflammation at the site of a central nervous system injury. The
CC   specification also details the identification of Beta2 integrin alphaD
CC   cDNAs and proteins, for use in raising the antibodies. Beta2 integrins
CC   are implicated in diseases such as LAD (leukocyte adhesion deficiency,
CC   inflammatory response, diabetes, multiple sclerosis, arthritis, graft
CC   atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative
CC   colitis, immune complex alveolitis and leukaemia. The present sequence is
CC   a Beta2 integrin alpha subunit sequence included for comparison with the
CC   Beta2 integrin alphaD protein sequences
XX
```

SQ Sequence 1153 AA;

Query Match 99.8%; Score 5862; DB 5; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||:|||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 5
AAO14428
ID AAO14428 standard; protein; 1153 AA.
XX
AC AAO14428;
XX
DT 03-MAY-2002 (first entry)
XX
DE Integrin Mac-1 alpha subunit.
XX
KW Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
KW open conformation; integrin related inflammatory disorder;
KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;
KW reperfusion; hypovolemic shock; infarction; cerebral shock;
KW viral infection; cancer; gene therapy; vaccine;
KW bioactive agent screening.
XX
OS Unidentified.
XX
PN WO200204521-A2.
XX
PD 17-JAN-2002.
XX
PF 09-JUL-2001; 2001WO-US021805.
XX
PR 07-JUL-2000; 2000US-0216600P.
XX
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
PA (BLOO-) CENT BLOOD RES.
XX
PI Springer T;
XX
DR WPI; 2002-148167/19.
XX
PT New integrin I domain protein having alteration in at least 2
PT noncontiguous regions and exits in an open conformation, useful for
PT treating, preventing or suppressing inflammatory or immunological
PT disorders.
XX
PS Example 1; Fig 1F; 90pp; English.
XX
CC The invention comprises structurally biased variant integrin inserted (I)
CC domain proteins, wherein the alterations to the protein occur in at least
CC two noncontinuous regions. Specifically the the variant integrin I domain
CC proteins are structurally biased to exist in the open conformation,
CC thereby altering the binding ability of the protein. The invention also
CC comprises nucleic acids encoding the variant integrin I domain proteins.
CC The integrin I domain proteins and nucleic acids are useful for treating,
CC preventing or suppressing integrin related inflammatory and immunological
CC disorders (e.g. rheumatoid arthritis). The variant integrin I domain
CC proteins and nucleic acids can also be used for treating: ischaemia/
CC reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral
CC infection; and cancer. The variant integrin I domain nucleic acids and
CC proteins may be used in gene therapy, as vaccines and to screen for
CC bioactive agents. The present amino acid sequence represents the Mac-1
CC alpha subunit of integrin
XX
SQ Sequence 1153 AA;

Query Match 99.8%; Score 5862; DB 5; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX
AC ADD25615;
XX
DT 15-JAN-2004 (first entry)
XX
DE Binding domain-immunoglobulin fusion protein-associated protein #85.
XX
KW Binding domain; immunoglobulin; fusion protein; cytostatic;
KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;
KW neuroprotective; hinge region; immunoglobulin heavy chain;
KW CH2 constant region; CH3 constant region; IgG1;
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.
XX
OS Unidentified.
XX
PN US2003118592-A1.
XX
PD 26-JUN-2003.
XX
PF 25-JUL-2002; 2002US-00207655.
XX
PR 17-JAN-2001; 2001US-0367358P.
PR 17-JAN-2002; 2002US-00053530.
PR 03-JUN-2002; 2002US-0385691P.
XX
PA (GENE-) GENECRAFT INC.
XX
PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
XX
DR WPI; 2003-801317/75.
XX
PT New binding domain-immunoglobulin fusion protein, useful for treating a
PT subject having or suspected of having a malignant condition or a B-cell
PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX
PS Disclosure; SEQ ID NO 176; 157pp; English.
XX
CC The invention relates to a binding domain-immunoglobulin fusion protein
CC comprising a binding domain polypeptide that is fused to an
CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
CC CH2 constant region polypeptide that is fused to the hinge region
CC polypeptide, and an immunoglobulin heavy chain CH3 constant region
CC polypeptide that is fused to the CH2 constant region polypeptide. The
CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin
CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge
CC region polypeptide, derived from (a) having 3 or more cysteine residues;
CC where the mutated human IgG1 immunoglobulin hinge region polypeptide
CC contains 2 cysteine residues, where the first cysteine is not mutated; a
CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
CC (a) having 3 or more cysteine residues, where the mutated human IgG1
CC immunoglobulin hinge region polypeptide contains no more than one
CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
CC polypeptide, derived from (a) having 3 or more cysteine residues; where
CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains
CC no cysteine residues. The binding domain-immunoglobulin fusion protein is
CC capable of at least one immunological activity comprising antibody
CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The
CC binding domain polypeptide is capable of specifically binding to an
CC antigen. Also included are an isolated polynucleotide encoding the
CC binding domain-immunoglobulin fusion protein, a recombinant expression
CC construct comprising the polynucleotide (operably linked to a promoter),
CC a host cell transformed or transfected with a recombinant expression
CC construct, producing the binding domain-immunoglobulin fusion protein, a

```
CC   pharmaceutical composition comprising the binding domain-immunoglobulin
CC   fusion protein or polynucleotide and a carrier, and treating a subject
CC   having or suspected of having a malignant condition or a B-cell disorder.
CC   The binding domain-immunoglobulin fusion protein is useful for treating a
CC   subject having or suspected of having a malignant condition or a B-cell
CC   disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,
CC   myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple
CC   sclerosis or autoimmune disease. The present sequence is a binding domain
CC   -immunoglobulin fusion protein-associated protein sequence. Note: The
CC   sequence data for this patent formed part of the printed specification
CC   and is also available in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not
CC   identified the sequences in the printed specification by their SEQ ID
CC   number therefore none of the sequences can be explicitly identified.
XX
SQ   Sequence 1153 AA;

  Query Match             99.8%;  Score 5862;  DB 7;  Length 1153;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||| |||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 7
AAR04136
ID   AAR04136 standard; protein; 1153 AA.
XX
AC   AAR04136;
XX
DT   09-SEP-2004  (revised)
DT   25-MAR-2003  (revised)
DT   07-SEP-1990  (first entry)
XX
DE   Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX
KW   Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
KW   non-specific defence system; integrin gene superfamily.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .16
FT                   /label= signal_peptide
FT   Modified-site   86. .88
FT                   /label= putative N-glycosylation site
FT   Modified-site   240. .242
FT                   /label= putative N-glycosylation site
FT   Modified-site   391. .393
FT                   /label= putative N-glycosylation site
FT   Modified-site   469. .471
FT                   /label= putative N-glycosylation site
FT   Modified-site   693. .695
FT                   /label= putative N-glycosylation site
FT   Modified-site   697. .699
FT                   /label= putative N-glycosylation site
FT   Modified-site   735. .737
FT                   /label= putative N-glycosylation site
FT   Modified-site   802. .804
FT                   /label= putative N-glycosylation site
FT   Modified-site   881. .883
FT                   /label= putative N-glycosylation site
FT   Modified-site   901. .903
FT                   /label= putative N-glycosylation site
FT   Modified-site   912. .914
FT                   /label= putative N-glycosylation site
FT   Modified-site   941. .943
FT                   /label= putative N-glycosylation site
FT   Modified-site   947. .949
FT                   /label= putative N-glycosylation site
FT   Modified-site   979. .981
FT                   /label= putative N-glycosylation site
```

```
FT   Modified-site    994. .996
FT                    /label= putative N-glycosylation site
FT   Modified-site    1022. .1024
FT                    /label= putative N-glycosylation site
FT   Modified-site    1045. .1047
FT                    /label= putative N-glycosylation site
FT   Modified-site    1051. .1053
FT                    /label= putative N-glycosylation site
FT   Modified-site    1076. .1078
FT                    /label= putative N-glycosylation site
FT   Region           1106. .1134
FT                    /label= putative_transmembrane_region
XX
PN   EP364690-A.
XX
PD   25-APR-1990.
XX
PF   17-AUG-1989;   89EP-00115159.
XX
PR   23-AUG-1988;   88US-00235353.
PR   09-MAR-1989;   89US-00321239.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Springer TA,  Corbi A;
XX
DR   WPI; 1990-125938/17.
DR   N-PSDB; AAQ04043.
XX
PT   New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating
PT   inflammation and viral infections, and in diagnosis.
XX
PS   Disclosure; Page ?; 3pp; English.
XX
CC   Mac-1 alpha subunit is involved in the response to inflammation, i.e.
CC   recognition of and migration to sites of inflammation. It also attaches
CC   to cellular substrates as part of this function making it useful in
CC   visualising endothelial tissue. Mac-1 is a member of the Integrin Gene
CC   superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25
CC   -MAR-2003 to correct PA field.)
CC
CC   Revised record issued on 09-SEP-2004 : Correction to feature table key
CC   and pages
XX
SQ   Sequence 1153 AA;

  Query Match             99.6%;  Score 5852;  DB 2;  Length 1153;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLLLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 8
ADM99589
ID   ADM99589 standard; protein; 1152 AA.
XX
AC   ADM99589;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Human integrin alphaM subunit precursor protein.
XX
KW   integrin alpha subunit; beta; antipsoriatic; thrombolytic; anticoagulant;
KW   osteopathic; cytostatic; immunosuppressive; antiinflammatory;
KW   neuroprotective; antisickling; immunotherapy; inflammatory;
KW   autoimmune disorder; thrombosis; cancer; osteoporosis;
KW   sickle cell anaemia; psoriasis; multiple sclerosis; human; precursor;
KW   alphaM.
XX
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   Misc-difference 965
FT                   /note= "Encoded by CCC"
XX
PN   WO2004007530-A2.
XX
PD   22-JAN-2004.
XX
PF   17-JUL-2003; 2003WO-US022301.
XX
PR   17-JUL-2002; 2002US-0396783P.
PR   17-JUL-2002; 2002US-0396790P.
PR   11-SEP-2002; 2002US-0410135P.
XX
PA   (BLOO-) CENT BLOOD RES INC.
XX
PI   Springer TA,  Takagi J;
XX
DR   WPI; 2004-122877/12.
DR   N-PSDB; ADM99588.
XX
PT   Novel modified integrin protein having extracellular domains of integrin
PT   alpha and beta subunits or integrin alphaII and beta3 subunit, useful for
PT   treating integrin mediated disorders.
XX
PS   Disclosure; SEQ ID NO 4; 232pp; English.
XX
CC   The invention relates to a novel isolated or recombinant modified
CC   integrin protein having extracellular domains of integrin alpha and beta
CC   subunits where one of the subunits has one or more mutations, an altered
CC   surface feature or an amino acid substitution or internal deletion,
CC   extracellular domains of the integrin beta subunit that comprise a
CC   mutation that alters a non-cysteine residue to cysteine or extracellular
CC   domains of integrin alpha and beta subunits. The polypeptide of the
CC   invention demonstrates antipsoriatic, thrombolytic, anticoagulant,
CC   osteopathic, cytostatic, immunosuppressive, antiinflammatory,
CC   neuroprotective and antisickling activities and may be useful for
CC   immunotherapy in order to prevent or treat an integrin-mediated disorder
CC   such as an inflammatory disorder, an autoimmune disorder, thrombosis,
CC   cancer, osteoporosis, sickle cell anaemia, psoriasis and multiple
CC   sclerosis. The current sequence is that of the human integrin alphaM
CC   subunit precursor protein of the invention.
XX
SQ   Sequence 1152 AA;

  Query Match             99.5%;  Score 5846.5;  DB 8;  Length 1152;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1128;  Conservative    7;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||||  |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 9
ADP12435
ID   ADP12435 standard; protein; 1152 AA.
XX
AC   ADP12435;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Protein encoded by mRNA of the invention #45.
XX
KW   transplant rejection; immune system; rheumatoid arthritis; lupus;
KW   inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX
OS   Homo sapiens.
XX
PN   WO2004042346-A2.
XX
PD   21-MAY-2004.
XX
PF   24-APR-2003; 2003WO-US012946.
```

XX
PR 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX
PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX
DR WPI; 2004-400724/37.
XX
PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
PS Claim 65; SEQ ID NO 2444; 1762pp; English.
XX
CC The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC protein that is encoded by the mRNA of the invention.
XX
SQ Sequence 1152 AA;

Query Match 99.4%; Score 5839.5; DB 8; Length 1152;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||:|||||||||||| |||:||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152
```

RESULT 10
ADP44061
ID ADP44061 standard; protein; 1163 AA.
XX
AC ADP44061;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human CD11C protein SEQ ID NO:14.
XX
KW HIV entry inhibitor; cell surface protein inhibitor; HIV infection;
KW anti-HIV; virucide; HIV uptake inhibitor; human; CD11C; integrin alpha X.
XX
OS Homo sapiens.
XX
PN WO2004053094-A2.
XX
PD 24-JUN-2004.
XX
PF 08-DEC-2003; 2003WO-US039208.
XX
PR 06-DEC-2002; 2002US-0431522P.
XX
PA (PPDP-) PPD DEV LP.
XX
PI Dunn SJ, Holzmayer TA;
XX
DR WPI; 2004-480928/45.
DR N-PSDB; ADP44060.
XX
PT Identifying an inhibitor of human immunodeficiency virus (HIV) entry into

PT a human host cell useful for preventing and/or treating HIV infection, by
PT identifying an inhibitor of a cell surface polypeptide such as CXCR-4.
XX
PS Claim 1; SEQ ID NO 14; 133pp; English.
XX
CC The present invention describes a method (M1) for identifying an
CC inhibitor of HIV entry into a human host cell. (M1) comprises identifying
CC an inhibitor of a cell surface polypeptide selected from CXCR-4 (352
CC amino acids, SEQ ID NO:8, ADP44055), CCR4 (360 amino acids, SEQ ID NO:10,
CC ADP44057), CCR7 (378 amino acids, SEQ ID NO:12, ADP44059), CD11C (1163
CC amino acids, SEQ ID NO:14, ADP44061), CD47 (323 amino acids, SEQ ID
CC NO:16, ADP44063), CD68 (354 amino acids, SEQ ID NO:18, ADP44065), CD69
CC (199 amino acids, SEQ ID NO:20, ADP44067), CD74 (566 amino acids, SEQ ID
CC NO:22, ADP44069), CSF3R (836 amino acids, SEQ ID NO:24, ADP44071), RARA
CC (462 amino acids, SEQ ID NO:26, ADP44073), GABBR1 (578 amino acids, SEQ
CC ID NO:28, ADP44075), P2X1 (150 amino acids, SEQ ID NO:30, ADP44077),
CC HELO1 (299 amino acids, SEQ ID NO:32, ADP44079), GPRK6 (576 amino acids,
CC SEQ ID NO:34, ADP44081), or PTK2B (1009 amino acids, SEQ ID NO:36,
CC ADP44083). Also described: (1) an inhibitor of a cell surface polypeptide
CC as described above in a human host cell preventing HIV entry into the
CC human host cell; (2) a pharmaceutical composition comprising an inhibitor
CC of (1) and a carrier; and (3) conferring resistance to HIV infection in
CC an individual, by administering the pharmaceutical composition of (2). An
CC inhibitor of HIV entry has anti-HIV and virucide activities, and can be
CC used as an HIV uptake inhibitor. (M1) is useful for identifying
CC protective compounds that inhibit entry of HIV into cells, useful for the
CC prevention and/or treatment of HIV infection. The present sequence
CC represents human integrin alpha X (CD11C), which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 1163 AA;

Query Match 59.3%; Score 3483; DB 8; Length 1163;
Best Local Similarity 61.2%; Pred. No. 1e-281;
Matches 691; Conservative 141; Mismatches 291; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||| |: :: ||| |||| | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               |||| |||||||||||:|||| |||||||||  |  | |: |||| |    | |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
               | : : ||::: || ||:|||||  :|  |:  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
               : ||||:||: : || ||:  : || | |:||| |: ||  ||: : |||::| ||||:|
Db        198 QTHFTFEEFRRSSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAAKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
               |||:| || | |:|||| ||  |:||| |||| ||::  | :||| :|||| ::|:|::
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
               :|:||| |||||:||||||||:| |||||| ||:||||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
               | |||      |||||:: : || ||:||||:  : |   |||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
               | :  |   | | |||||:||||||||||||:|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDSDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
               ||| | || |||||||||| |||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
               || | |||||||||||:|| |||||||:||||||| |||||:|||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
               ||| |   |:| | |: |: ||| :|||  :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
               ||||| ||   || | |||| :  : :|||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
               :||| || || ||||:|| |||| |||  |||||||  |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAADAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
               | |  | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
               |    |||| ||| ||   :::||  ||||  || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
               |||||||||||| ||:||   ||||||: | |  | |  |:|||:|||||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
               | |||  :|   :|:  :| |  | :::  |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
               | | :||| : |||||||| | :   : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
               | :| ::| || |||||||:|||||||||||| |||:||||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 11
ADN02004
ID ADN02004 standard; protein; 1163 AA.
XX
AC ADN02004;
XX
DT 17-JUN-2004 (first entry)
XX
DE Human inflammatory bowel disease marker - CD11c protein.
XX
KW marker gene; inflammatory bowel disease; IBD; FcgammaR IIIa;
KW FcgammaR IIIb; Mig; NRG-2; hexokinase; HM74; REG III; LPAP; Mip-1(beta);
KW L-selectin; EGFL6; IDO; IL-8; CD11c; TLR2; ulcerative colitis;
KW Crohn's disease.
XX
OS Homo sapiens.
XX
PN JP2004065120-A.
XX
PD 04-MAR-2004.
XX
PF 07-AUG-2002; 2002JP-00229705.
XX
PR 07-AUG-2002; 2002JP-00229705.
XX
PA (SUMU) SUMITOMO SEIYAKU KK.
XX
DR WPI; 2004-209124/20.
DR N-PSDB; ADN01990.
XX
PT Novel marker gene of inflammatory bowel disease (IBD) comprising base
PT sequence of e.g., Fc gamma R IIIa, Fc gamma R IIIb, Mig, NRG-2,
PT hexokinase 3, HM74, CD11c, TLR2 gene, useful for screening therapeutic

```
PT   agent for IBD.
XX
PS   Example; SEQ ID NO 27; 151pp; Japanese.
XX
CC   The invention comprises marker genes for inflammatory bowel disease (IBD)
CC   - FcgammaR IIIa, FcgammaR IIIb, Mig, NRG-2, hexokinase, HM74, REG III,
CC   LPAP, Mip-1(beta), L-selectin, EGFL6, IDO, IL-8, CD11c, and TLR2 genes.
CC   The DNA and protein sequences of the invention are useful in the
CC   detection and treatment of IBD (e.g. ulcerative colitis and Crohn's
CC   disease). The present amino acid sequence represents an IBD marker
CC   protein of the invention.
XX
SQ   Sequence 1163 AA;

  Query Match             59.0%;  Score 3469;  DB 8;  Length 1163;
  Best Local Similarity   61.0%;  Pred. No. 1.5e-280;
  Matches  689;  Conservative  141;  Mismatches  293;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||   | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||||||||||:|||| ||||||||||   |  | |: |||||  |    |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || ||:||||||   :|     |:  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              : ||||:||:   || ||:  : || | |:||| |: ||  ||: : |||::| ||||:|
Db        198 QTHFTFEEFRRTSNPLSLLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:| || | |:|||| ||  |:||| |||| ||::  | :||| :|||| ::|:|::
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:|||||||||:| |||||| ||:||||||   | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |   |||||||||||| |    :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:|||||||||||::|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| ||||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  |||||||||||:|| |||||||:|||||||| |||||| |||:| ||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :|||   :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :|||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:||  ||| |||  |||||| |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::||  ||||  || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATPDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| ||:||  ||||||: | |  | |  |:|||:||||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | ||| :|   :|: :| | | ::: | ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :      : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:|||||||||| |||:||||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143


RESULT 12
ADQ17510
ID   ADQ17510 standard; protein; 1163 AA.
XX
AC   ADQ17510;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated protein - SEQ ID 327.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 327; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   protein of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 1163 AA;
```

```
Query Match             59.0%;  Score 3469;  DB 8;  Length 1163;
Best Local Similarity   61.0%;  Pred. No. 1.5e-280;
Matches  689;  Conservative  141;  Mismatches  293;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||    | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||||||||||:|||| ||||||||||  |  | |: |||||  |    |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || ||:||||||    :|      |:  |: | ::   | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              : ||||:||:   || ||:  : || | |:||| |: ||  ||: : |||::| ||||:|
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:| || | |:|||| ||  |:||| |||| ||::  | :||| :|||| ::|:|::
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:|||||||:| |||||| ||:|||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||       |||||:: : || |:||||:  : |   |||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:||||||||||||::|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  ||||||||||||:|| |||||||:||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :|||  :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              |||||  ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:||  ||| |||  |||||||  |:||||:|  |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::||  ||||  || ||:||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| ||:||   ||||||: | |  | |  |:|||:|||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| | | :::  |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
                | | :||| : |||||||| |  :     : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:||||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 13
AAR07120
ID AAR07120 standard; protein; 1163 AA.
XX
AC AAR07120;
XX
DT 25-MAR-2003 (revised)
DT 05-FEB-1991 (first entry)
XX
DE p150.95 alpha subunit encoded by clone lambdaX47.
XX
KW p150,95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;
KW rhinovirus.
XX
OS Synthetic.
XX

FH	Key	Location/Qualifiers
FT	Region	1. .19
FT		/label= signal peptide
FT	Region	20. .44
FT		/label= N-terminus
FT	Modified-site	61. .63
FT		/label= glycosylation site
FT	Modified-site	89. .91
FT		/label= glycosylation site
FT	Modified-site	385. .387
FT		/label= glycosylation site
FT	Modified-site	392. .394
FT		/label= glycosylation site
FT	Modified-site	697. .699
FT		/label= glycosylation site
FT	Modified-site	735. .737
FT		/label= glycosylation site
FT	Modified-site	899. .901
FT		/label= glycosylation site
FT	Modified-site	904. .906
FT		/label= glycosylation site
FT	Modified-site	939. .941
FT		/label= glycosylation site
FT	Modified-site	1050. .1052
FT		/label= glycosylation site
FT	Domain	1108. .1133
FT		/label= transmembrane

XX
PN WO9010646-A.
XX
PD 20-SEP-1990.
XX
PF 09-MAR-1990; 90WO-US001257.
XX
PR 09-MAR-1990; 90WO-US001257.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Corbi AA, Springer TA;
XX
DR WPI; 1990-304985/40.
DR N-PSDB; AAQ06068.
XX
PT Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-
PT unit of p150.95 cell surface adhesion receptor, opt. together with a beta
PT chain of CD-18 family.
XX
PS Disclosure; Fig 3; 59pp; English.
XX

```
CC   Clone lambda X47 was isolated from a cDNA library constructed from total
CC   RNA extracted from phorbol myristate acetate stimulated HL-60
CC   myelomonocytic cells. The library was screened with oligonucleotide
CC   probes based on tryptic peptide fragments of p150.95. The sequence can be
CC   attached to appropriate control elements and expressed in prokaryotic and
CC   eukaryotic cells. The protein can be used to treat or prevent rhinoviral
CC   infection because it interacts with ICAM-1 and inhibits cell-virus
CC   attachment. It can also be used as an anti-inflammatory agent. See also
CC   AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003
CC   to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1163 AA;

  Query Match             58.9%;  Score 3460;  DB 2;  Length 1163;
  Best Local Similarity   60.9%;  Pred. No. 8.7e-280;
  Matches  687;  Conservative  141;  Mismatches  295;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||   | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               |||| ||||||||||||:|||| ||||||||||   |  | |: ||||| |    |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || ||:|||||   :|     |:  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              : ||||:||:   || ||:  : || | |:||| |: ||  ||: : |||::| ||||:|
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:| || | |:|||| ||  |:||| |||| ||::  | :||| :|||| ::|:|::
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:|||||||||:| |||||| ||:|||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |   |||||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:|||||||||||||::|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| ||||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  ||||||||||||:|| |||||||:||||||| ||||||  |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :|||  :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:||  ||| |||  |||||||  |:||||:| |:||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSA--PVA 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::||  |||| || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| ||:||   ||||||: | |  | |  |:|||:|||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| |  | :::  |  |||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSQKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :     : :|| ||| |: ||:: ||||  |:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQERFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:| |||||||||  |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGALLLLALITAVLYKVGFFKRQYKEMMEE 1143


RESULT 14
AAW65091
ID   AAW65091 standard; protein; 1163 AA.
XX
AC   AAW65091;
XX
DT   28-SEP-1998  (first entry)
XX
DE   Human Beta-integrin CD11c subunit protein.
XX
KW   Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW   type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
KW   lung inflammation; acute respiratory distress syndrome; CD11c subunit;
KW   rheumatoid arthritis.
XX
OS   Homo sapiens.
XX
PN   US5728533-A.
XX
PD   17-MAR-1998.
XX
PF   07-JUN-1995;   95US-00485618.
XX
PR   23-DEC-1993;   93US-00173497.
PR   05-AUG-1994;   94US-00286889.
PR   21-DEC-1994;   94US-00362652.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Van Der Vieren M,  Gallatin WM;
XX
DR   WPI; 1998-206565/18.
XX
PT   Screening assay for modulators of integrin binding - using immobilised or
PT   labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.
XX
PS   Example 5; Fig 1A-D; 106pp; English.
XX
CC   This sequence represents a human beta-integrin CD11c subunit which is
CC   used to describe a method for identifying compounds that modulate the
CC   interaction of the beta-integrin alpha-d subunit with a binding partner
CC   of alpha-d which involves contacting an alpha-d polypeptide with an alpha
CC   -d binding partner, one of which is immobilised and the other of which is
CC   labelled, in the presence of a test compound, and determining if the
CC   compound affects binding between the alpha-d polypeptide and alpha-d
CC   binding partner, where the alpha-d polypeptide is alpha-d or its fragment
CC   comprising the cytoplasmic, transmembrane or extracellular domain of
CC   alpha-d. Compounds that modulate alpha-d binding could be used to treat
CC   diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC   asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC   and rheumatoid arthritis
XX
```

SQ Sequence 1163 AA;

Query Match 58.6%; Score 3446; DB 2; Length 1163;
Best Local Similarity 60.8%; Pred. No. 1.3e-278;
Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| |||    | |||||||:|:|||| | |||| ||||:|||| 
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||:|||| |||||||||   | | |: ||||| |   |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || ||:|||||   :|    |:  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              : ||||:||:   || ||:  : || | |:||| |: ||   ||: : |||::| ||||:|
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAIKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:| || | |:|||| ||  |:||| |||| ||::  | :||| :|||| ::|:|::
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:|||||||||:| |||||| ||:|||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |   ||||||||||||||   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |    | | |||||:|||||||||||||::|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVIGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  ||||||||||||:|| |||||||:|||||||| |||||| ||||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :|||  :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||  ||   ||:| |||| :  : :|||| ||  | ||:|:|| | ||:|||||
Db        677 LDLALAPGRLSPRAIFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || |||:||  ||| |||  ||||||| |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |  ||  | |   |:
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTC--CSAPVG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::||  |||| ||   ::||| |||:||||:|||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              ||||||||||:||:||   ||||||: | |  | |  |:|||:|||| ||:|: | ||
Db        915 QLELPVKYAVYIVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| | |  | :::  |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : ||||||||| |  :      : :|| |||:|: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEIIFDTSVYSQLPGQEAFMRAQTI 1094

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| ||:|||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPIPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 15
AAB07361
ID AAB07361 standard; protein; 1163 AA.
XX
AC AAB07361;
XX
DT 17-JAN-2001 (first entry)
XX
DE Human CD11c protein sequence.
XX
KW Human; macrophage infiltration inhibition; alpha_d integrin;
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW rheumatoid arthritis; central nervous system injury; CD11c.
XX
OS Homo sapiens.
XX
PN WO200029446-A1.
XX
PD 25-MAY-2000.
XX
PF 16-NOV-1999; 99WO-US027139.
XX
PR 16-NOV-1998; 98US-00193043.
PR 08-JUL-1999; 99US-00350259.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gallatin MW, Van Der Vieren M;
XX
DR WPI; 2000-387751/33.
XX
PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous system
PT injury sites.
XX
PS Example 5; Fig 1; 270pp; English.
XX
CC Integrins are a class of membrane-associated molecules that participate
CC in cellular adhesion. Integrins are made up of an alpha subunit and a
CC beta subunit. One class of human integrins are restricted to expression
CC in white blood cells and have a common beta2 subunit: the leukocyte
CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC have an important role in immune and inflammatory responses. The present
CC protein sequence is the human integrin alpha subunit CD11c. This sequence
CC was used in an alignment to identify a novel beta2 integrin alpha
CC subunit: alpha_d (AAA60014 and AAB07359). The present sequence has
CC approximately 66% identity to the protein sequence of alpha_d. The
CC Alpha_d gene and protein may be useful in therapy for diseases linked to
CC aberrant alpha_d function e.g. Type I diabetes, atherosclerosis, multiple
CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
CC (LAD). In addition, anti-alpha_d monoclonal antibodies may be used in the
CC inhibition of macrophage infiltration at the site of a central nervous
CC system injury. The monoclonal antibodies can also be used to detect and
CC diagnose Crohn's disease
XX
SQ Sequence 1163 AA;

Query Match 58.6%; Score 3446; DB 3; Length 1163;
Best Local Similarity 60.8%; Pred. No. 1.3e-278;

Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAIKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        318 DFDALKDIQNQLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        438 QVSRQWRMKAEVIGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALAPGRLSPRAIFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTC--CSAPVG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYIVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEIIFDTSVYSQLPGQEAFMRAQTI 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPIPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

Search completed: November 9, 2004, 12:08:24
Job time : 108.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 12:17:49 ; Search time 97.5 Seconds
(without alignments)
4119.157 Million cell updates/sec

Title: US-09-902-481B-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5876	100.0	1137	10	US-09-902-481A-5	Sequence 5, Appli
2	5862	99.8	1153	9	US-09-350-259-3	Sequence 3, Appli
3	5862	99.8	1153	10	US-09-902-481A-1	Sequence 1, Appli
4	5862	99.8	1153	10	US-09-891-943-3	Sequence 3, Appli
5	5862	99.8	1153	14	US-10-144-259-30	Sequence 30, Appl
6	5862	99.8	1153	14	US-10-207-655-176	Sequence 176, App
7	5855	99.6	1137	10	US-09-902-481A-6	Sequence 6, Appli
8	5851	99.6	1137	10	US-09-902-481A-4	Sequence 4, Appli
9	5846.5	99.5	1152	9	US-09-945-265-4	Sequence 4, Appli
10	5836	99.3	1137	10	US-09-902-481A-3	Sequence 3, Appli
11	3469	59.0	1163	14	US-10-116-275-204	Sequence 204, App
12	3446	58.6	1163	9	US-09-350-259-4	Sequence 4, Appli
13	3446	58.6	1163	10	US-09-891-943-4	Sequence 4, Appli
14	3411	58.0	1161	9	US-09-350-259-2	Sequence 2, Appli
15	3411	58.0	1161	10	US-09-891-943-2	Sequence 2, Appli
16	3395.5	57.8	1161	9	US-09-350-259-99	Sequence 99, Appl
17	3395.5	57.8	1161	10	US-09-891-943-99	Sequence 99, Appl
18	3232.5	55.0	1161	9	US-09-350-259-55	Sequence 55, Appl
19	3232.5	55.0	1161	10	US-09-891-943-55	Sequence 55, Appl
20	3222	54.8	1161	9	US-09-350-259-53	Sequence 53, Appl
21	3222	54.8	1161	10	US-09-891-943-53	Sequence 53, Appl
22	3211.5	54.7	1151	9	US-09-350-259-37	Sequence 37, Appl
23	3211.5	54.7	1151	10	US-09-891-943-37	Sequence 37, Appl
24	3201	54.5	1155	9	US-09-350-259-46	Sequence 46, Appl
25	3201	54.5	1155	10	US-09-891-943-46	Sequence 46, Appl
26	1848	31.4	369	13	US-10-087-192-1212	Sequence 1212, Ap
27	1557.5	26.5	1170	9	US-09-945-265-2	Sequence 2, Appli
28	1553.5	26.4	1170	15	US-10-261-164-1	Sequence 1, Appli
29	1530	26.0	1223	16	US-10-408-765A-295	Sequence 295, App
30	1359.5	23.1	1086	16	US-10-408-765A-1871	Sequence 1871, Ap
31	1229.5	20.9	494	9	US-09-350-259-103	Sequence 103, App
32	1229.5	20.9	494	10	US-09-891-943-103	Sequence 103, App
33	1153	19.6	1179	14	US-10-177-550-2	Sequence 2, Appli
34	1153	19.6	1179	14	US-10-173-551-2	Sequence 2, Appli
35	1151.5	19.6	413	9	US-09-350-259-101	Sequence 101, App
36	1151.5	19.6	413	10	US-09-891-943-101	Sequence 101, App
37	1103.5	18.8	1151	10	US-09-984-130-103	Sequence 103, App
38	1103.5	18.8	1151	10	US-09-836-353A-103	Sequence 103, App
39	1103.5	18.8	1179	10	US-09-918-715-250	Sequence 250, App
40	1103.5	18.8	1179	17	US-10-474-794-250	Sequence 250, App
41	1101	18.7	1188	14	US-10-291-265-810	Sequence 810, App
42	1097	18.7	1188	14	US-10-291-265-338	Sequence 338, App
43	1093.5	18.6	1189	10	US-09-984-130-35	Sequence 35, Appl
44	1093.5	18.6	1189	10	US-09-836-353A-35	Sequence 35, Appl
45	1093.5	18.6	1189	15	US-10-262-839-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-902-481A-5
; Sequence 5, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-5

Query Match 100.0%; Score 5876; DB 10; Length 1137;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
```

```
Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

RESULT 2
US-09-350-259-3
; Sequence 3, Application US/09350259

; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.8%; Score 5862; DB 9; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676
```

```
Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 3
US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:
US-09-902-481A-1

Query Match 99.8%; Score 5862; DB 10; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||| |||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 4
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1

```
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT:  Van der Vieren, Monica
;  TITLE OF INVENTION: No. US20030077278A1el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/891,943
;  CURRENT FILING DATE:  2001-06-26
;  PRIOR APPLICATION NUMBER: 09/193,043
;  PRIOR FILING DATE: 1998-11-16
;  PRIOR APPLICATION NUMBER: 08/286,889
;  PRIOR FILING DATE: 1994-08-05
;  PRIOR APPLICATION NUMBER: 08/362,652
;  PRIOR FILING DATE: 1994-12-21
;  PRIOR APPLICATION NUMBER: 08/943,363
;  PRIOR FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-891-943-3

  Query Match             99.8%;  Score 5862;  DB 10;  Length 1153;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 5
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
;  APPLICANT: Arnaout, M. Amin
;  APPLICANT:  Li, Rui
;  APPLICANT:  Xiong, Jian-Ping
;  TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: 00786-548001
;  CURRENT APPLICATION NUMBER: US/10/144,259
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: US 09/758,493
;  PRIOR FILING DATE: 2001-01-11
;  PRIOR APPLICATION NUMBER: US 60/221,950
;  PRIOR FILING DATE: 2000-07-31
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-144-259-30

  Query Match             99.8%;  Score 5862;  DB 14;  Length 1153;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||::|
```

```
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153

RESULT 6
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
;  APPLICANT: Ledbetter, Jeffrey A.
;  APPLICANT:  Hayden-Ledbetter, Martha S.
;  TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
;  FILE REFERENCE: 390069.401C1
;  CURRENT APPLICATION NUMBER: US/10/207,655
;  CURRENT FILING DATE:  2002-07-25
;  NUMBER OF SEQ ID NOS: 426
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-207-655-176

  Query Match             99.8%;  Score 5862;  DB 14;  Length 1153;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180

Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240

Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916
```

```
Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 7
US-09-902-481A-6
; Sequence 6, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6

Query Match 99.6%; Score 5855; DB 10; Length 1137;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
Db        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
Db        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

RESULT 8
US-09-902-481A-4
; Sequence 4, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic

US-09-902-481A-4

Query Match 99.6%; Score 5851; DB 10; Length 1137;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||:|||||||||||| ||::|||||||||||||||||||||||
Db        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||| |:
Db        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

RESULT 9
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; TITLE OF INVENTION: DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: CBN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4

Query Match 99.5%; Score 5846.5; DB 9; Length 1152;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 1; Indels 1; Gaps 1;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||:|||||||||||| |||:|||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555
```

```
Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152
```

RESULT 10
US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 99.3%; Score 5836; DB 10; Length 1137;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

RESULT 11
US-10-116-275-204
; Sequence 204, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 204
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-204

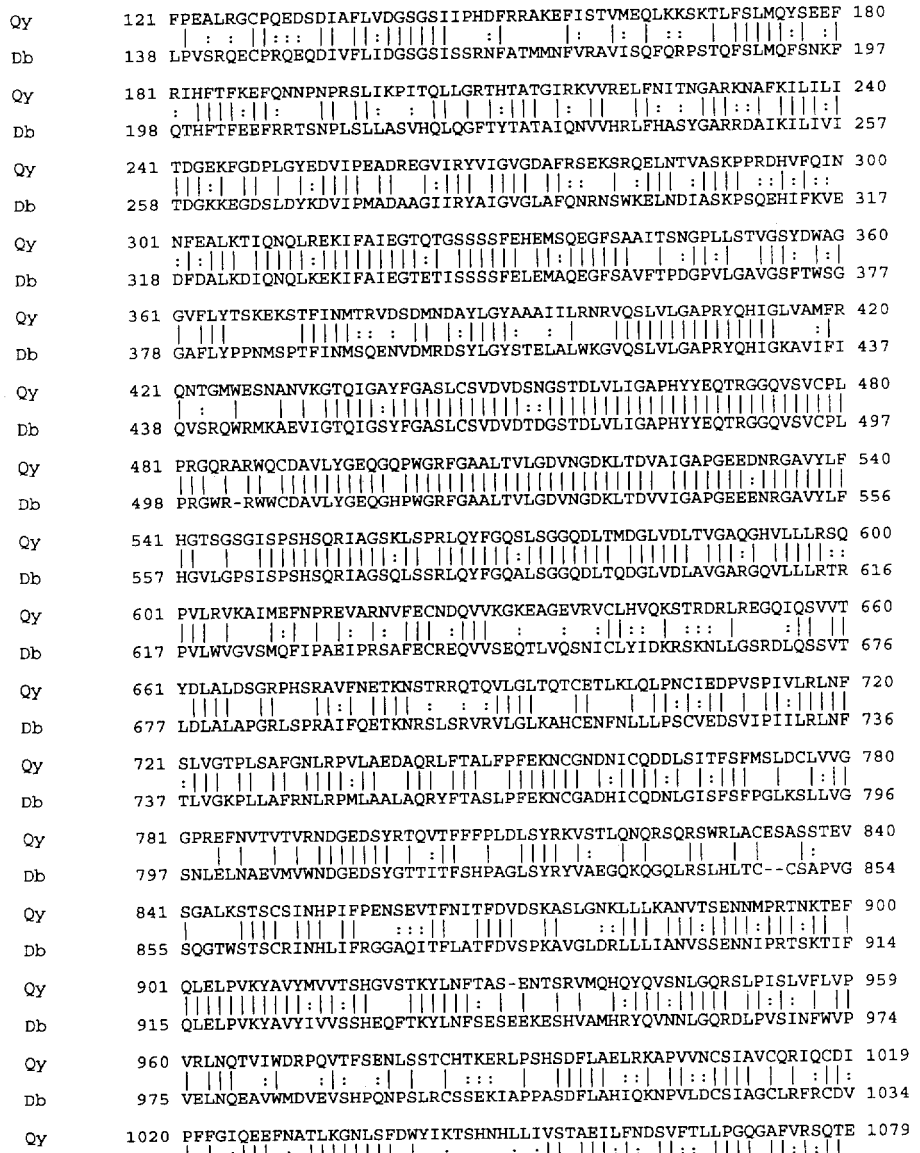
Query Match 59.0%; Score 3469; DB 14; Length 1163;
Best Local Similarity 61.0%; Pred. No. 2e-285;
Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||| |: :: ||| |||| | ||||||:| |||| | |||| ||||:|||| 
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||:|||| |||||||||| | | |: ||||| |  | |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || ||:|||||  :|     |: |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              : ||||:||:   || ||:  : || | |:||| |: ||  ||: : |||::| ||||:|
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:| || | |:|||| ||  |:||| |||| ||::  | :||| :|||| ::|:|::
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:|||||||||:| |||||| ||:||||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||       |||||:: : || |:||||:  : |   |||||||||||||  |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:||||||||||||::|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||| |||||||:|||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  |||||||||||:|| |||||||:|||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :||| :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :|||| ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:||  ||| |||  ||||||| |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |  |||| ||| ||   :::||  |||| || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||||| ||:||  ||||||: | |  | |  |:|||:|||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| |  | ::: |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :     : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 12
US-09-350-259-4
; Sequence 4, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-4

Query Match 58.6%; Score 3446; DB 9; Length 1163;
Best Local Similarity 60.8%; Pred. No. 1.8e-283;
Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;

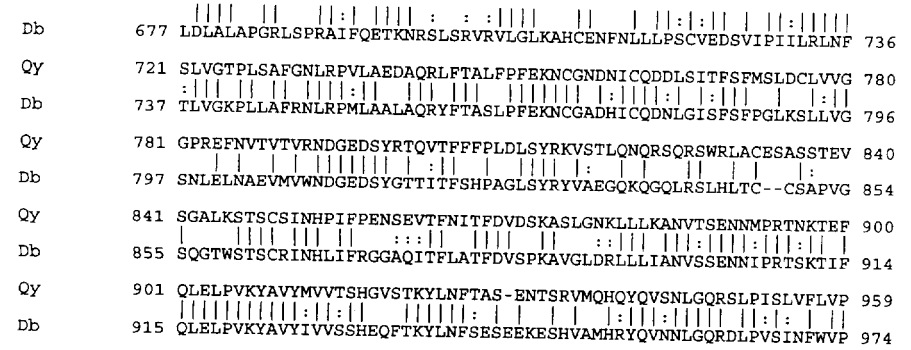
```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||| |: :: ||| |||| | |||||||:|:|||| | |||| ||||:|||| 
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||:|||| |||||||||| | | |: ||||| |  | |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137
```



RESULT 13
US-09-891-943-4
; Sequence 4, Application US/09891943

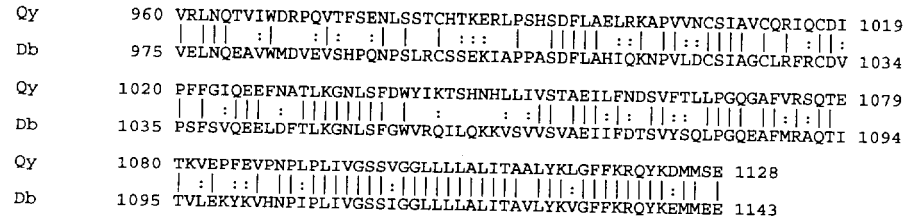
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-4

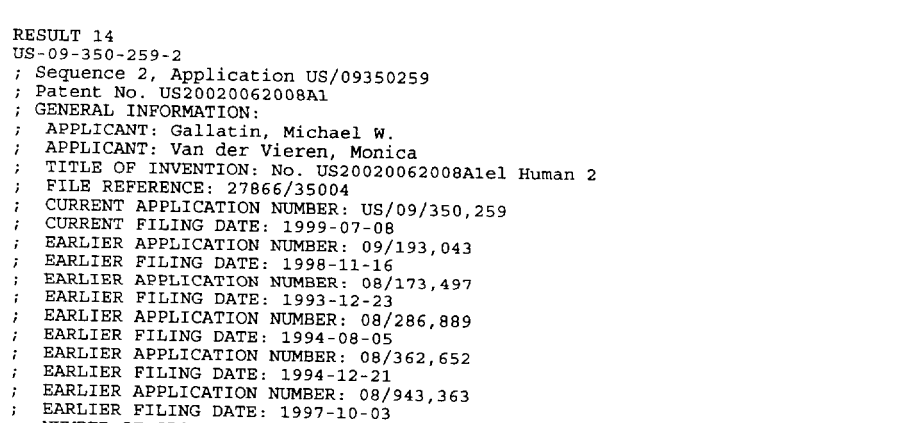
Query Match 58.6%; Score 3446; DB 10; Length 1163;
Best Local Similarity 60.8%; Pred. No. 1.8e-283;
Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;

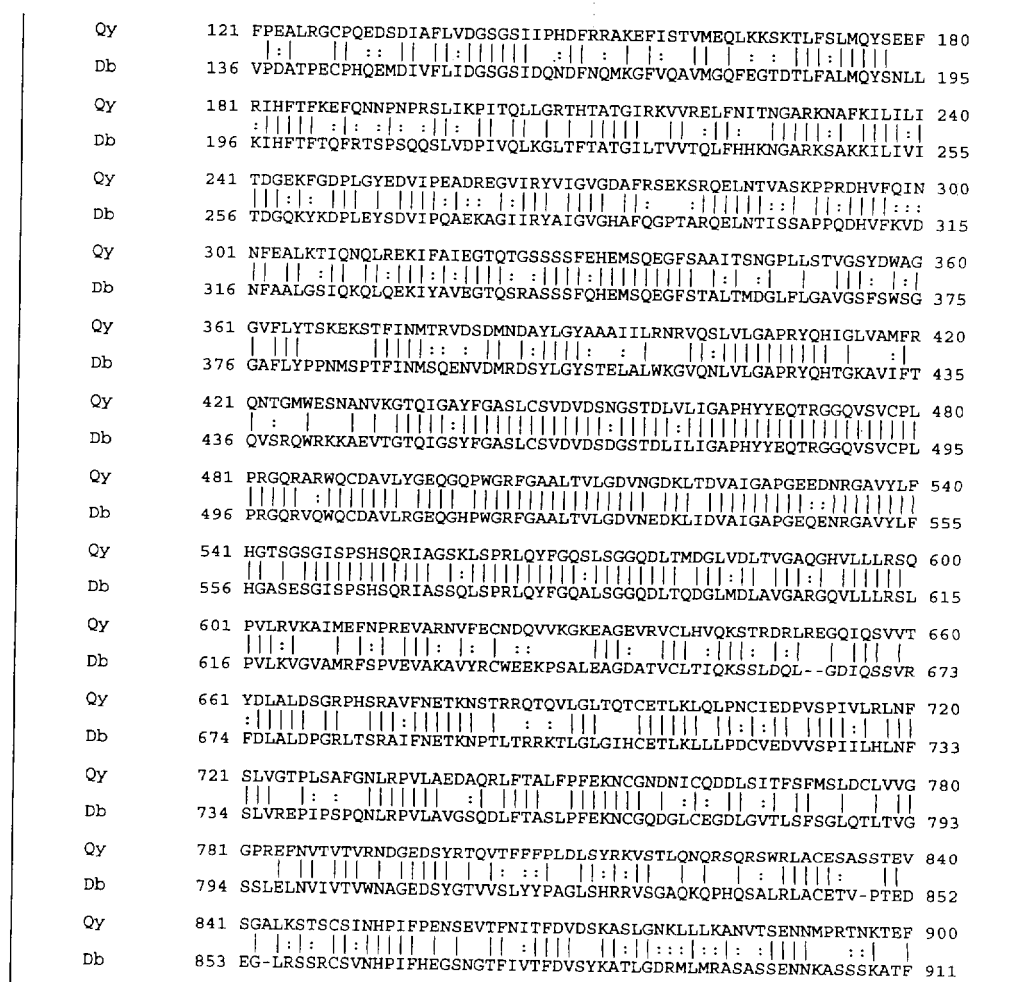


RESULT 14
US-09-350-259-2
; Sequence 2, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-2

Query Match 58.0%; Score 3411; DB 9; Length 1161;
Best Local Similarity 59.4%; Pred. No. 1.7e-280;
Matches 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;







RESULT 15
US-09-891-943-2
; Sequence 2, Application US/09891943
; Publication No. US20030077278A1

; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-2

Query Match 58.0%; Score 3411; DB 10; Length 1161;
Best Local Similarity 59.4%; Pred. No. 1.7e-280;
Matches 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    |||:| ||||||||  |||:||||| |:||||| | || |  :|| |:||
Db         17 FNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQTGRLYDCAAATGMCQPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               | :  |||||||||:|||:|:  :|||||||:|: | ||:| || | | ||   :  |
Db         77 PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-EIIQT 135

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              |:|   || :: || ||:||||||  :|| : | |:  || | : : |||:|||||
Db        136 VPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFEGTDTLFALMQYSNLL 195

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              :||||| :|: :|: :||: || || | | ||||| || :||:  |||||:| ||||:|
Db        196 KIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHHKNGARKSAKKILIVI 255

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||:|: ||| | ||||:|:: |:||| |||| ||:   :|||||::| ||:||||::
Db        256 TDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELNTISSAPPQDHVFKVD 315

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              || || :|| ||:|||:|:|||: :||||:||||||||| |:| :|  |  |||: |:|
Db        316 NFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGSFSWSG 375

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||     |||||:: : || |:||||:  : |   ||:|||||||| |   :|
Db        376 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVLGAPRYQHTGKAVIFT 435

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:|||||||||||||||:|||||:||||||||||||||||||||
Db        436 QVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPHYYEQTRGGQVSVCPL 495

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||| :||||||| |||| ||||||||||||||||| ||| |||||||||::||||||||
Db        496 PRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAIGAPGEQENRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || | |||||||||||| |:||||||||||:|||||||| |||:|| |||:| ||||||
Db        556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLLRSL 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |||:|   | |:| |||: |: | ::     |||:  ||| :|||: |:|  | |||  |
Db        616 PVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              :||||| ||  |||:|||||| |  : : |||   |||||| ||:|:|| ||||:| |||
Db        674 FDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCVEDVVSPIILHLNF 733

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||  |: :  |||||||  :| ||||  |||||||  | :|: || :| ||  |  | ||
Db        734 SLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDLGVTLSFSGLQTLTVG 793

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | || ||| | ||||| | |: ::|  ||:|:||  | |  | : |||||:   ||
Db        794 SSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              | |:|: ||:|||||| | |  || :|||| ||:||:::|::|: :||||   ::|  |
Db        853 EG-LRSSRCSVNHPIFHEGSNGTFIVTFDVSYKATLGDRMLMRASASSENNKASSSKATF 911

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNF-TASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| :::    |||| || |: |   :  :|:|:|:|| || | ||: | ||
Db        912 QLELPVKYAVYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQRDLAISINFWVP 971

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | ||   :||      |::|   | :: : | ||||| :: ::|:::|||| | : :||:
Db        972 VLLNGVAVWDVVMEAPSQSL--PCVSERKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1029

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :|    :|:|| ||| |: ||:: |||| ||:|:| |
Db       1030 PSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQLPGQEAFMRAQME 1089

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
                :|  || | :|:|:||||| ||||||||| |||||||| ||:|: :
Db       1090 MVLEEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRHYKEMLED 1138
```

Search completed: November 9, 2004, 12:46:58
Job time : 102.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:57 ; Search time 23 Seconds
(without alignments)
3278.416 Million cell updates/sec

Title: US-09-902-481B-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5862	99.8	1153	1	US-08-173-497-3	Sequence 3, Appli
2	5862	99.8	1153	1	US-08-286-889-3	Sequence 3, Appli
3	5862	99.8	1153	1	US-08-485-618-3	Sequence 3, Appli
4	5862	99.8	1153	1	US-08-362-652-3	Sequence 3, Appli
5	5862	99.8	1153	2	US-08-605-672-3	Sequence 3, Appli
6	5862	99.8	1153	2	US-08-482-293A-3	Sequence 3, Appli
7	5862	99.8	1153	2	US-08-943-363-3	Sequence 3, Appli
8	5862	99.8	1153	3	US-09-193-043-3	Sequence 3, Appli
9	5862	99.8	1153	4	US-09-688-307A-3	Sequence 3, Appli
10	5862	99.8	1153	4	US-09-350-259-3	Sequence 3, Appli
11	5831.5	99.2	1152	2	US-08-476-062A-43	Sequence 43, Appl
12	5831.5	99.2	1152	5	PCT-US96-01314-43	Sequence 43, Appl
13	5831.5	99.2	1152	6	5424399-2	Patent No. 5424399
14	3469	59.0	1163	2	US-08-476-062A-44	Sequence 44, Appl
15	3469	59.0	1163	5	PCT-US96-01314-44	Sequence 44, Appl
16	3446	58.6	1163	1	US-08-173-497-4	Sequence 4, Appli
17	3446	58.6	1163	1	US-08-286-889-4	Sequence 4, Appli
18	3446	58.6	1163	1	US-08-485-618-4	Sequence 4, Appli
19	3446	58.6	1163	1	US-08-362-652-4	Sequence 4, Appli
20	3446	58.6	1163	2	US-08-605-672-4	Sequence 4, Appli
21	3446	58.6	1163	2	US-08-482-293A-4	Sequence 4, Appli
22	3446	58.6	1163	2	US-08-943-363-4	Sequence 4, Appli
23	3446	58.6	1163	3	US-09-193-043-4	Sequence 4, Appli
24	3446	58.6	1163	4	US-09-688-307A-4	Sequence 4, Appli
25	3446	58.6	1163	4	US-09-350-259-4	Sequence 4, Appli
26	3411	58.0	1161	1	US-08-173-497-2	Sequence 2, Appli
27	3411	58.0	1161	1	US-08-286-889-2	Sequence 2, Appli
28	3411	58.0	1161	1	US-08-485-618-2	Sequence 2, Appli
29	3411	58.0	1161	1	US-08-362-652-2	Sequence 2, Appli
30	3411	58.0	1161	2	US-08-605-672-2	Sequence 2, Appli
31	3411	58.0	1161	2	US-08-482-293A-2	Sequence 2, Appli
32	3411	58.0	1161	2	US-08-943-363-2	Sequence 2, Appli
33	3411	58.0	1161	3	US-09-193-043-2	Sequence 2, Appli
34	3411	58.0	1161	4	US-09-688-307A-2	Sequence 2, Appli
35	3411	58.0	1161	4	US-09-350-259-2	Sequence 2, Appli
36	3395.5	57.8	1161	1	US-08-485-618-99	Sequence 99, Appl
37	3395.5	57.8	1161	2	US-08-605-672-99	Sequence 99, Appl
38	3395.5	57.8	1161	2	US-08-482-293A-99	Sequence 99, Appl
39	3395.5	57.8	1161	2	US-08-943-363-99	Sequence 99, Appl
40	3395.5	57.8	1161	3	US-09-193-043-99	Sequence 99, Appl
41	3395.5	57.8	1161	4	US-09-688-307A-99	Sequence 99, Appl
42	3395.5	57.8	1161	4	US-09-350-259-99	Sequence 99, Appl
43	3232.5	55.0	1161	3	US-09-193-043-55	Sequence 55, Appl
44	3232.5	55.0	1161	4	US-09-688-307A-55	Sequence 55, Appl
45	3232.5	55.0	1161	4	US-09-350-259-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

```
Qy        61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy       121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
             ||||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||
Db       137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy       181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db       197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy       241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||:|
Db       257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy       301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy       361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy       421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy       481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy       541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy       601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy       661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy       721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy       781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy       841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy       901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy       961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy      1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy      1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1097 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Mich
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5470953el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,889
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  P38,659
;      REFERENCE/DOCKET NUMBER:  27866/32168
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-286-889-3

  Query Match             99.8%;  Score 5862;  DB 1;  Length 1153;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376
```

```
Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 3
US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5728533el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,618
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32797
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-485-618-3

  Query Match             99.8%;  Score 5862;  DB 1;  Length 1153;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||| |||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||:|||||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||:|||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 4
US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-362-652-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 5
US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-672-3

Query Match 99.8%; Score 5862; DB 2; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180

Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240

Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036
```

```
Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 6
US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5831029el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,293A
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32684
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-482-293A-3

  Query Match             99.8%;  Score 5862;  DB 2;  Length 1153;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5837478el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/943,363
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32684
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-943-363-3

  Query Match             99.8%;  Score 5862;  DB 2;  Length 1153;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153

RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT: Van der Vieren, Monica
;  TITLE OF INVENTION: No. 6251395el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/193,043
;  CURRENT FILING DATE: 1998-11-16
;  EARLIER APPLICATION NUMBER: 08/173,497
;  EARLIER FILING DATE: 1993-12-23
;  EARLIER APPLICATION NUMBER: 08/286,889
;  EARLIER FILING DATE: 1994-08-05
;  EARLIER APPLICATION NUMBER: 08/362,652
;  EARLIER FILING DATE: 1994-12-21
;  EARLIER APPLICATION NUMBER: 08/943,363
;  EARLIER FILING DATE: 1997-10-03
```

```
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;    LENGTH: 1153
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-193-043-3

  Query Match             99.8%;  Score 5862;  DB 3;  Length 1153;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 9
US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT:  Van der Vieren, Monica
;  TITLE OF INVENTION: No. 6432404el Human Beta-2
;  FILE REFERENCE: 27866/36646
;  CURRENT APPLICATION NUMBER: US/09/688,307A
;  CURRENT FILING DATE:  2000-10-13
;  PRIOR APPLICATION NUMBER: 09/193,043
;  PRIOR FILING DATE: 1998-11-16
;  PRIOR APPLICATION NUMBER: 08/605,672
;  PRIOR FILING DATE: 1996-02-22
;  PRIOR APPLICATION NUMBER: 08/173,497
;  PRIOR FILING DATE: 1993-12-23
;  PRIOR APPLICATION NUMBER: 08/286,889
;  PRIOR FILING DATE: 1994-08-05
;  PRIOR APPLICATION NUMBER: 08/362,652
;  PRIOR FILING DATE: 1994-12-21
;  PRIOR APPLICATION NUMBER: 08/943,363
;  PRIOR FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;    LENGTH: 1153
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-688-307A-3

  Query Match             99.8%;  Score 5862;  DB 4;  Length 1153;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376
Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436
Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496
Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616
Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676
Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736
Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796
Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856
Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916
Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976
Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036
Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096
Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153

RESULT 10
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT: Van der Vieren, Monica
;  TITLE OF INVENTION: No. 6620915el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/350,259
;  CURRENT FILING DATE: 1999-07-08
;  EARLIER APPLICATION NUMBER: 09/193,043
;  EARLIER FILING DATE: 1998-11-16
;  EARLIER APPLICATION NUMBER: 08/173,497
;  EARLIER FILING DATE: 1993-12-23
;  EARLIER APPLICATION NUMBER: 08/286,889
;  EARLIER FILING DATE: 1994-08-05
;  EARLIER APPLICATION NUMBER: 08/362,652
;  EARLIER FILING DATE: 1994-12-21
;  EARLIER APPLICATION NUMBER: 08/943,363
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;  LENGTH: 1153
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-09-350-259-3

  Query Match             99.8%;  Score 5862;  DB 4;  Length 1153;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1129;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76
Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136
Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196
Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256
Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316
Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376
Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436
Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496
Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616
Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676
Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736
Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796
Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856
Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
```

```
Db        857 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

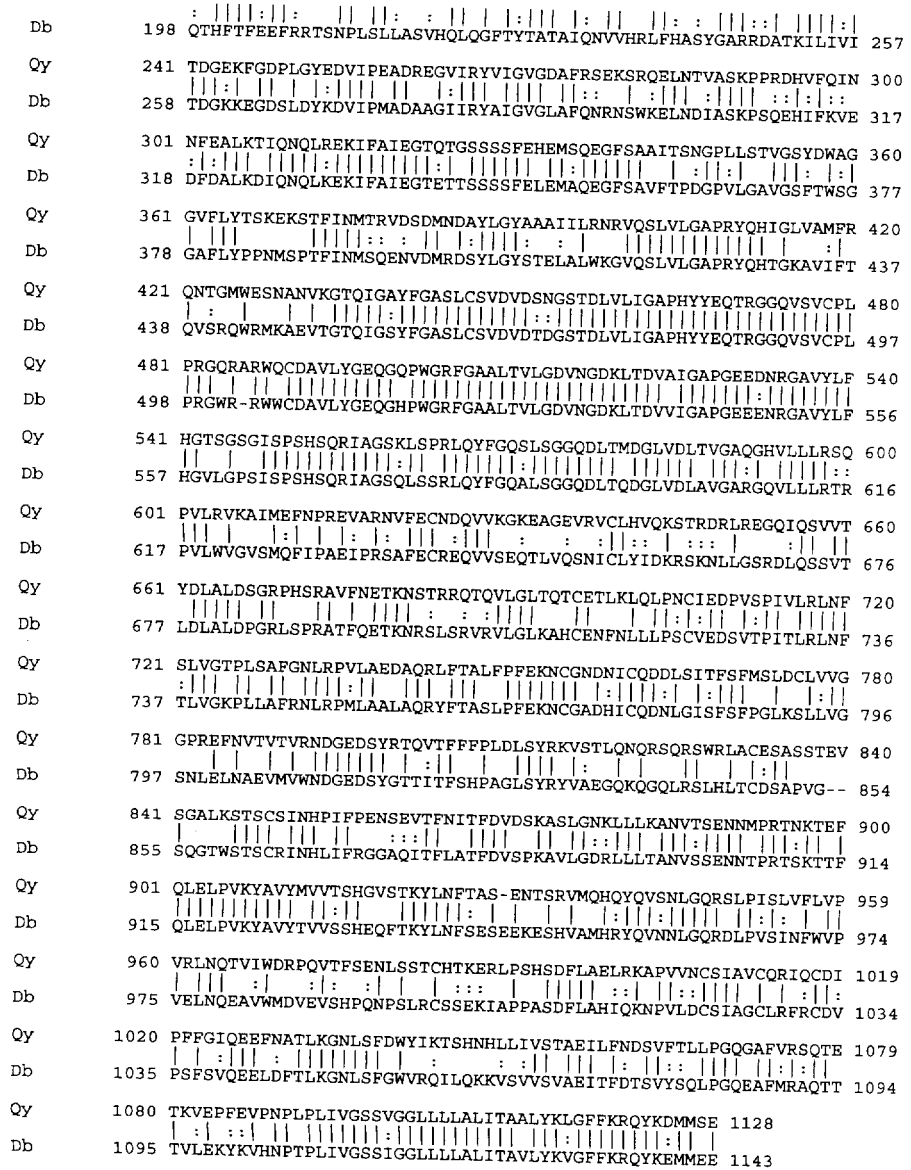
```
RESULT 11
US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
;  GENERAL INFORMATION:
;    APPLICANT:  Arnaout, M. Amin
;    TITLE OF INVENTION:  CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
;    TITLE OF INVENTION:  RESPONSES WITH BETA2 INTEGRINS
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,062A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/216,081
;      FILING DATE:  21-MAR-1994
;      APPLICATION NUMBER:  07/637,830
;      FILING DATE:  04-JAN-1991
;      APPLICATION NUMBER:  07/539,842
;      FILING DATE:  18-JUN-1990
;      APPLICATION NUMBER:  07/212,573
;      FILING DATE:  28-JUN-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Freeman, John W.
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  00786/068003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1152 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-476-062A-43

  Query Match           99.2%;  Score 5831.5;  DB 2;  Length 1152;
  Best Local Similarity  99.0%;  Pred. No. 0;
  Matches 1126;  Conservative    7;  Mismatches    3;  Indels    1;  Gaps    1;
```

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180

Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240

Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        796 GPRESNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

```
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 12
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
;  GENERAL INFORMATION:
;    APPLICANT:  M. Amin Arnaout
;    TITLE OF INVENTION:  METHODS FOR IDENTIFYING INTEGRIN
;    TITLE OF INVENTION:  ANTAGONISTS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/01314
;      FILING DATE:  30-JAN-96
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/380,167
;      FILING DATE:  30-JAN-95
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  John W. Freeman
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  00786/267001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1152
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
PCT-US96-01314-43

  Query Match           99.2%;  Score 5831.5;  DB 5;  Length 1152;
  Best Local Similarity  99.0%;  Pred. No. 0;
  Matches 1126;  Conservative    7;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPRESNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 13
5424399-2
;Patent No. 5424399
;    APPLICANT: ARNAOUT, M. AMIN
;    TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
;    NUMBER OF SEQUENCES: 12
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/78,871
;      FILING DATE: 16-JUN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 539,842
;      FILING DATE: 18-JUN-1990
;      APPLICATION NUMBER: 212,573
;      FILING DATE: 28-JUN-1988
;SEQ ID NO:2:
;      LENGTH: 1152
5424399-2

  Query Match           99.2%;  Score 5831.5;  DB 6;  Length 1152;
```

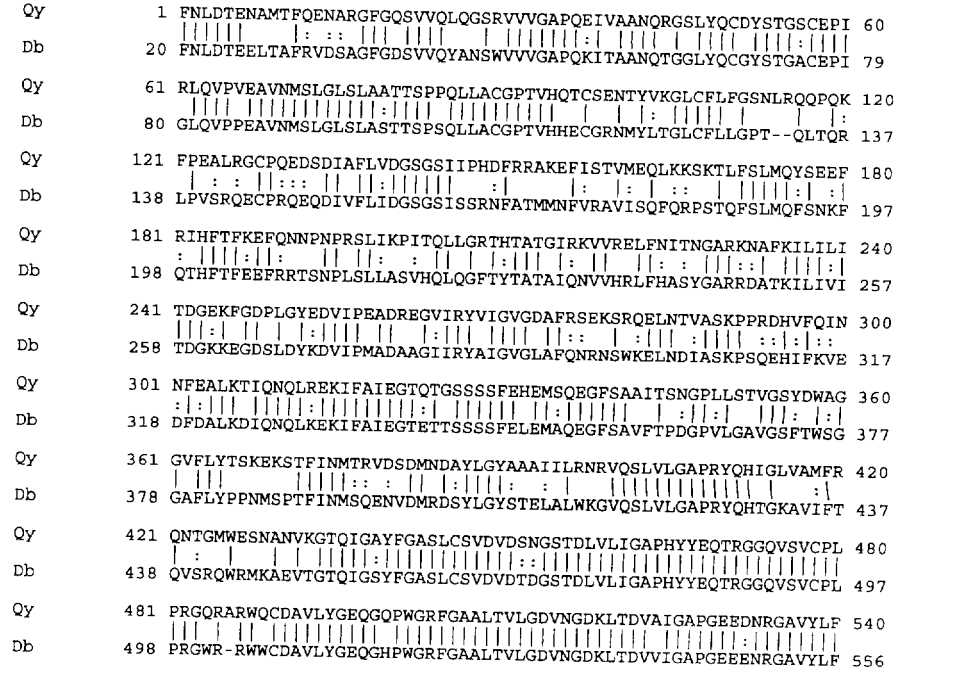
```
Best Local Similarity   99.0%;  Pred. No. 0;
Matches 1126;  Conservative    7;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||:|||||||||||||| |||:||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||::|
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||:|
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPRESNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152
```

```
RESULT 14
US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
;  GENERAL INFORMATION:
;    APPLICANT:  Arnaout, M. Amin
;    TITLE OF INVENTION:  CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
;    TITLE OF INVENTION:  RESPONSES WITH BETA2 INTEGRINS
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,062A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/216,081
;      FILING DATE:  21-MAR-1994
;      APPLICATION NUMBER:  07/637,830
;      FILING DATE:  04-JAN-1991
;      APPLICATION NUMBER:  07/539,842
;      FILING DATE:  18-JUN-1990
;      APPLICATION NUMBER:  07/212,573
;      FILING DATE:  28-JUN-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Freeman, John W.
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  00786/068003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1163 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-476-062A-44

  Query Match           59.0%;  Score 3469;  DB 2;  Length 1163;
  Best Local Similarity   61.0%;  Pred. No. 7e-284;
  Matches  689;  Conservative  141;  Mismatches  293;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||   | ||||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||||||||||:|||| |||||||||  |  | |: ||||| |     |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || ||:||||||   :|     |:  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
```



RESULT 15
PCT-US96-01314-44
; Sequence 44, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-44

Query Match 59.0%; Score 3469; DB 5; Length 1163;
Best Local Similarity 61.0%; Pred. No. 7e-284;
Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;



```
Qy       541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db       557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy       601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db       617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy       661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db       677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy       721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db       737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy       781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db       797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy       841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db       855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy       901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db       915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy       960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db       975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

Search completed: November 9, 2004, 12:21:09
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:37 ; Search time 27.25 Seconds
(without alignments)
4014.622 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5868	99.9	1153	1	RWHU1B	cell surface glyco
2	4476	76.2	1153	2	S00551	leukocyte surface
3	3483	59.3	1163	1	RWHU1C	cell surface glyco
4	1548.5	26.4	1170	2	S03308	cell surface glyco
5	1533.5	26.1	1163	2	I56126	lymphocyte fuction
6	1149	19.6	1179	2	A53213	integrin alpha-E c
7	1102.5	18.8	1151	2	A45226	integrin alpha-1 c
8	1084	18.5	1170	2	I45914	integrin alpha 2 s
9	1072	18.2	1178	2	S44142	VLA-2 protein homo
10	1069	18.2	1181	2	A33998	integrin alpha-2 c
11	1060	18.0	1180	2	A35854	integrin alpha-1 c
12	663	11.3	1039	2	A41131	lymphocyte-Peyer's
13	633	10.8	1038	2	S06046	integrin alpha-4 c
14	630	10.7	1035	2	I58409	integrin alpha-9 c
15	614.5	10.5	1041	2	T31437	integrin alpha cha
16	579.5	9.9	1054	2	JC7294	alphaP integrin -
17	572.5	9.7	1051	2	A35761	cell surface glyco
18	567.5	9.7	1053	2	I55534	VLA-3 alpha subuni
19	555.5	9.5	1053	2	S44250	integrin alpha-5 c
20	543.5	9.3	1034	2	A36108	integrin alpha-V c
21	535	9.1	1044	2	T10050	integrin alpha-v c
22	532	9.1	1049	2	A27079	fibronectin recept
23	532	9.1	1073	2	B36429	integrin alpha-6 c
24	530.5	9.0	1072	2	A38457	integrin alpha-6 c
25	529.5	9.0	1051	2	A40021	integrin alpha-3 c
26	526	9.0	1048	2	A27421	integrin alpha-5 c
27	525.5	8.9	1091	2	A41543	integrin alpha-6 c
28	517	8.8	1044	2	S16516	integrin alpha-8 c
29	496.5	8.5	1146	2	S40311	integrin - fruit f
30	495.5	8.4	1039	2	A34269	integrin alpha-2b
31	494	8.4	1394	2	A29637	position-specific
32	490	8.3	1137	2	JC5950	integrin alpha-7 c
33	488	8.3	1135	2	I61186	alpha-7 integrin -
34	486	8.3	126	2	B30892	leukocyte adhesion
35	485.5	8.3	1037	2	A60163	glycoprotein IIb -
36	469.5	8.0	1106	2	S38783	integrin alpha cha
37	466	7.9	1226	2	S44824	F54F2.1 protein -
38	454	7.7	1045	2	S60571	integrin alpha v c
39	445.5	7.6	1139	2	S28277	hypothetical prote
40	422	7.2	1115	2	T09403	integrin alpha cha
41	417.5	7.1	1115	2	T09433	integrin alpha cha
42	391	6.7	764	2	I36916	glycoprotein IIb -
43	309	5.3	1086	2	T18523	integrin alpha cha
44	299	5.1	604	2	I36917	glycoprotein IIb -
45	296.5	5.0	272	2	A55348	integrin alpha-1 -

ALIGNMENTS

RESULT 1
RWHU1B
cell surface glycoprotein CD11b precursor [validated] - human
N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Ma
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567
R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.
J. Biol. Chem. 263, 12403-12411, 1988
A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD
B.
A;Reference number: A31108; MUID:88315033; PMID:2457584
A;Accession: A31108
A;Molecule type: mRNA
A;Residues: 1-1153 <COR>
A;Cross-references: UNIPROT:P11215; GB:J03925; NID:g187284; PIDN:AAA59544.1; PID:g307148
A;Note: part of this sequence was confirmed by protein sequencing
R;Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
J. Cell Biol. 106, 2153-2158, 1988
A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor M
A;Reference number: A28915; MUID:88257215; PMID:2454931
A;Accession: A28915
A;Molecule type: mRNA
A;Residues: 1-499,501-965,'P',967-1153 <ARN>
A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AAA594
A;Note: the authors translated the codon TAC for residue 1129 as Thr
A;Note: part of this sequence, including the amino end of the mature protein, was confir
R;Shelley, C.S.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991
A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg
A;Reference number: A41600; MUID:92073318; PMID:1683702
A;Accession: A41600
A;Molecule type: DNA
A;Residues: 1-9 <SHE>
A;Cross-references: GB:M76724; NID:g180018; PIDN:AAA58410.1; PID:g553215
R;Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesi
A;Reference number: A94193; MUID:88190151; PMID:2833753
A;Accession: A30892
A;Molecule type: mRNA
A;Residues: 917-1042 <AR2>
A;Cross-references: GB:M18044
R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
A;Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor
A;Reference number: A32218; MUID:89098893; PMID:2563162
A;Accession: A32218
A;Molecule type: mRNA
A;Residues: 9-1153 <HIC>
A;Cross-references: GB:J04145; NID:g189068; PIDN:AAA59903.1; PID:g386975

A;Note: part of this sequence was confirmed by protein sequencing
R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A;Reference number: A46526; MUID:93123748; PMID:8419480
A;Accession: A46526
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-499,501-1153 <FLE>
A;Cross-references: GB:S52227; NID:g263047; PIDN:AAB24821.1; PID:g263049
A;Note: the last three bases of intron 13, CAG, are included in some but not all mature
A;Note: sequence extracted from NCBI backbone (NCBIP:121963)
R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A;Title: N-terminal sequence of human leukocyte glycoprotein Mo1: conservation across sp
A;Reference number: A90664; MUID:87076671; PMID:3539202
A;Accession: A26091
A;Molecule type: protein
A;Residues: 17-31 <PIE>
A;Experimental source: granulocytes
R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A;Title: Characterization of the myeloid-specific CD11b promoter.
A;Reference number: I52567; MUID:92144986; PMID:1346576
A;Accession: I52567
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-9 <RES>
A;Cross-references: GB:M84477; NID:g180184; PIDN:AAA51960.1; PID:g553219
C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C;Genetics:
A;Gene: GDB:ITGAM; CR3A
A;Cross-references: GDB:120599; OMIM:120980
A;Map position: 16p11.2-16p11.2
A;Note: promoter contains a GATA motif and two Sp1 consensus binding sites
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F;17-1108/Domain: extracellular #status predicted <EXT>
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;465-473/Region: calcium/magnesium binding #status predicted
F;530-538/Region: calcium/magnesium binding #status predicted
F;593-601/Region: calcium/magnesium binding #status predicted
F;1109-1134/Domain: transmembrane #status predicted <TMM>
F;1135-1153/Domain: intracellular #status predicted <INT>
F;86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 2
S00551
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
N;Alternate names: complement-3 receptor alpha chain
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S00551; I59078
R;Pytela, R.
EMBO J. 7, 1371-1378, 1988
A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the
A;Reference number: S00551; MUID:88312584; PMID:3044779
A;Accession: S00551
A;Molecule type: DNA
A;Residues: 1-1153 <PYT>
A;Cross-references: UNIPROT:P05555; EMBL:X07640; NID:g52982; PIDN:CAA30479.1; PID:g5298
A;Note: the authors translated the codon CAC for residue 569 as Gln

```
R;Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts,
Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A;Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
A;Reference number: I59078; MUID:86287312; PMID:2942940
A;Accession: I59078
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 11-44 <RES>
A;Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
C;Genetics:
A;Gene: Mac-1
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C;Keywords: cell adhesion; glycoprotein; transmembrane protein
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1106-1129/Domain: transmembrane #status predicted <TMM>

  Query Match           76.2%;  Score 4476;  DB 2;  Length 1153;
  Best Local Similarity  74.3%;  Pred. No. 6.1e-300;
  Matches  845;  Conservative  142;  Mismatches  149;  Indels    2;  Gaps    2;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||: |||||||:||||:|||| |: ||| |||| | || |:||||||| | ||
Db         17 FNLDTEHPMTFQENAKGFGQNVVQLGGTSVVVAAPQEAKAVNQTGALYQCDYSTSRCHPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               |||| |||||||||||| :| | |||||||||||| | ||||| |||:|||||| : ||:
Db         77 PLQVPPEAVNMSLGLSLAVSTVPQQLLACGPTVHQNCKENTYVNGLCYLFGSNLLRPPQQ 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              |||||| |||::||| |||||||||    ||::|||||||||| ||||||||||||:||
Db        137 FPEALRECPQQESDIVFLIDGSGSINNIDFQKMKEFVSTVMEQFKKSKTLFSLMQYSDEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |||||| :|: ||:||| | || || ||| ||:|:|||:||| : |||||:|| |||:||
Db        197 RIHFTFNDFKRNPSPRSHVSPIKQLNGRTKTASGIRKVVRELFHKTNGARENAAKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              |||||||||| |:|||||||| ||||||||||:||   :||:||:|||||  :|||||:
Db        257 TDGEKFGDPLDYKDVIPEADRAGVIRYVIGVGNAFNKPQSRRELDTIASKPAGEHVFQVD 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||| ||||||:|||||||||||||:|||||||||||||:|||||||| :|||:||||
Db        317 NFEALNTIQNQLQEKIFAIEGTQTGSTSSFEHEMSQEGFSASITSNGPLLGSVGSFDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||||:| |||| |||||||||||||||:|:|||||||||||||||||||| |||
Db        377 GAFLYTSKDKVTFINTTRVDSDMNDAYLGYASAVILRNRVQSLVLGAPRYQHIGLVVMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              :| | || : ::||:|||:|||||||||:|::|:|:|:|||||||:|||||||||||
Db        437 ENFGTWEPHTSIKGSQIGSYFGASLCSVDMDADGNTNLILIGAPHYYEKTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||:|:|:|:|| ||||||||||||||||||||||||||||::|:||||:|
Db        497 PRG-RARWQCEALLHGDQGHPWGRFGAALTVLGDVNGDKLTDVAIGAPGEQENQGAVYIF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              :| | : :| \|| || |:  || |||||||||||:|||||||:|| ||||||:||||:|
Db        556 YGASIASLSASHSHRIIGAHFSPGLQYFGQSLSGGKDLTMDGLMDLAVGAQGHLLLLRAQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||::| |||:|::|||:|| | :||:| |:|||||||| |:|:|:|||||| ||| ||
Db        616 PVLRLEATMEFSPKKVARSVFACQEQVLKNKDAGEVRVCLRVRKNTKDRLREGDIQSTVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              |||||| |   || |:||||:|||:||| || | |||||| ||:|::| ||||:||||:
Db        676 YDLALDPVRSRIRAFFDETKNNTRRRTQVFGLMQKCETLKLILPDCVDDSVSPIILRLNY 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || :||||||||| |||| |||:||||||||||:||||||||| | | || ||||
Db        736 TLVGEPLRSFGNLRPVLAMDAQRFFTAMFPFEKNCGNDSICQDDLSITMSAMGLDTLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRL-ACESASSTE 839
              ||::||::||:|||||||| |||| ::|  ||||| |  ||  ::: | :  ||:||:|
Db        796 GPQDFNMSVTLRNDGEDSYGTQVTVYYPSGLSYRKDSASQNPLTKKPWFVKPAESSSSSE 855

Qy        840 VSGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTE 899
              ||||||: :||||||| |||||||:|||||| || |||||||| | ||||| ||:||:
Db        856 GHGALKSTTWNINHPIFPPANSEVTFNVTFDVDSHASFGNKLLLKAIVASENNMSRTHKTK 915

Qy        900 FQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              ||||||||||:||:|||  | :|||||||| ||:|:|||||| :|||||||||:|:|| :|
Db        916 FQLELPVKYAIYMIVTSDESSIRYLNFTASEMTSKVIQHQYQFNNLGQRSLPVSVVFWIP 975

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              |::|   :|| ||| ||:|||| |||::: | ||:|  :| : ||:|||:|||:|||||:
Db        976 VQINNVTVWDHPQVIFSQNLSSACHTEQKSPPHSNFRDQLERTPVLNCSVAVCKRIQCDL 1035

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | || || |||||||||||||||| |||:||: ||||||| | |||||  ::|||:||
Db       1036 PSFNTQEIFNVTLKGNLSFDWYIKTSHGHLLLVSSTEILFNDSAFALLPGQESYVRSKTE 1095

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||:|| ||:|||||||:|||:||||||| |||||||||||||:|  |  | ||
Db       1096 TKVEPYEVHNPVPLIVGSSIGGLVLLALITAGLYKLGFFKRQYKDMMNEAAPQDAPPQ 1153


RESULT 3
RWHU1C
cell surface glycoprotein CD11c precursor - human
N;Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A36584; A35543; S00864
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A;Reference number: A36584
A;Contents: erratum
A;Accession: A36584
A;Molecule type: DNA
A;Residues: 1-1163 <COR>
A;Cross-references: UNIPROT:P20702
A;Note: this revision to the sequence from reference A35543 includes the carboxyl end
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule
A;Reference number: A35543; MUID:90153906; PMID:2303426
A;Accession: A35543
A;Molecule type: DNA
A;Residues: 1-834 <CO2>
A;Note: this sequence has been revised in reference A36584
R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A;Reference number: S00864; MUID:88166645; PMID:3327687
A;Accession: S00864
A;Molecule type: mRNA
A;Residues: 1-755,'L',757-1163 <CO3>
A;Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830
A;Note: part of this sequence was confirmed by protein sequencing
C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my
C;Genetics:
A;Gene: GDB:ITGAX; CD11C
A;Cross-references: GDB:119758; OMIM:151510
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C;Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F;20-1107/Domain: extracellular #status predicted <EXT>
F;149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F;1108-1133/Domain: transmembrane #status predicted <TMM>
F;1134-1163/Domain: intracellular #status predicted <INT>
F;61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre
```

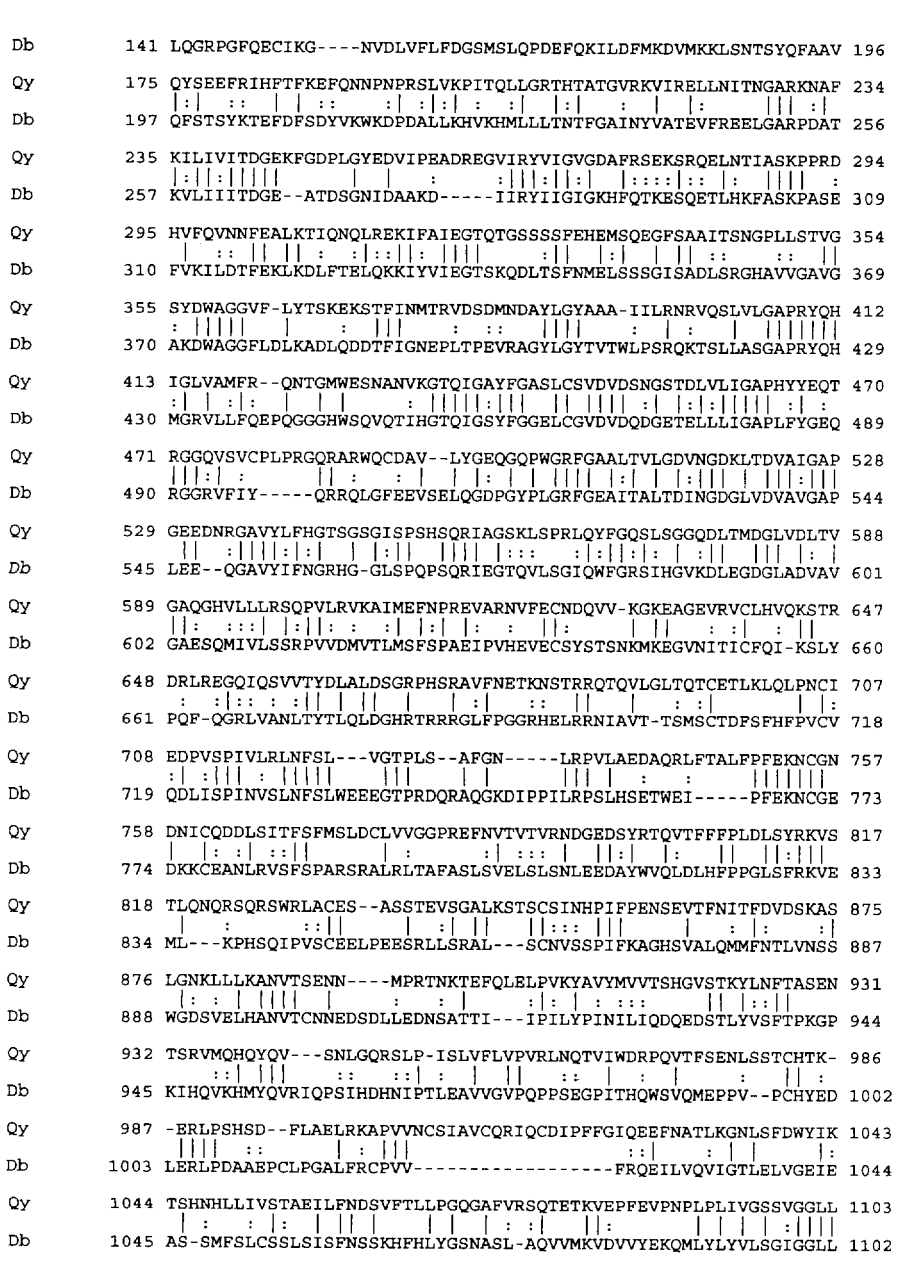
Query Match 59.3%; Score 3483; DB 1; Length 1163;
Best Local Similarity 61.4%; Pred. No. 1.8e-231;
Matches 693; Conservative 138; Mismatches 292; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAADAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 4
S03308
cell surface glycoprotein CD11a precursor - human
N;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte functio
C;Species: Homo sapiens (man)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: S03308; A47458; A47565; A48759; S36044
R;Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.
J. Cell Biol. 108, 703-712, 1989
A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha subuni
A;Reference number: S03308; MUID:89139587; PMID:2537322
A;Accession: S03308
A;Molecule type: mRNA
A;Residues: 1-1170 <LAR>
A;Cross-references: UNIPROT:P20701; UNIPROT:Q9UBC8; EMBL:Y00796; NID:g31421; PIDN:CAA68
A;Note: part of this sequence was confirmed by protein sequencing
R;Cornwell, R.D.; Gollahon, K.A.; Hickstein, D.D.
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993
A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a) pr
A;Reference number: A47458; MUID:93248261; PMID:8097887
A;Accession: A47458
A;Molecule type: DNA
A;Residues: 1-20 <COR>
A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBIP:130863)
R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993
A;Title: Identification of cell-specific and developmentally regulated nuclear factors
A;Reference number: A47565; MUID:93281759; PMID:8099450
A;Accession: A47565
A;Molecule type: DNA
A;Residues: 1-20 <SHE>
A;Cross-references: GB:M95609
R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.
J. Biol. Chem. 268, 19305-19311, 1993
A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter.
A;Reference number: A48759; MUID:93374910; PMID:8103515
A;Accession: A48759
A;Molecule type: DNA
A;Residues: 1-20 <NUE>
A;Cross-references: EMBL:Z22804; NID:g311405; PIDN:CAA80461.1; PID:g311406
C;Genetics:
A;Gene: GDB:ITGAL; CD11A
A;Cross-references: GDB:119757; OMIM:153370
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen; tr
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted
F;154-317/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 26.4%; Score 1548.5; DB 2; Length 1170;
Best Local Similarity 34.5%; Pred. No. 4.3e-98;
Matches 405; Conservative 208; Mismatches 460; Indels 101; Gaps 37;

```
Qy          1 FNLDTENAMTFQ--ENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58
Db         26 YNLDVRGARSFSPPRAGRHFGYRVLQV-GNGVIVGAPGE---GNSTGSLYQCQSGTGHCL 81

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLR--- 115
Db         82 PVTLR-GSNYTSKYLGMTLATDPTDGSILACDPGLSRTCDQNTYLSGLCYLFRQNLQGPM 140

Qy        116 -QQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLM 174
```



RESULT 5
I56126

lymphocyte fuction-associated molecule-1-alpha - mouse
C;Species: Mus musculus (house mouse)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: I56126
R;Kaufmann, Y.; Tseng, E.; Springer, T.A.
J. Immunol. 147, 369-374, 1991
A;Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit
A;Reference number: I56126; MUID:91268576; PMID:2051027
A;Accession: I56126
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1163 <RES>
A;Cross-references: UNIPROT:P24063; GB:M60778; NID:g198785; PIDN:AAA39426.1; PID:g19878
C;Genetics:
A;Gene: LFA-1
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
F;151-315/Domain: von Willebrand factor type A repeat homology <VWA1>

Query Match 26.1%; Score 1533.5; DB 2; Length 1163;
Best Local Similarity 34.0%; Pred. No. 4.6e-97;
Matches 398; Conservative 217; Mismatches 458; Indels 99; Gaps 36;

```
Qy        763 DDLSITFSFMSLDCLVVGGP------REFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKV 816
              :|:::   |      ||         |  |: | |||:|  ::   ||  ||:|||
Db        775 ANLTLSSPARS-------GPLRLMSSASLAVEWTLSNSGEDAYWVRLDLDFPRGLSFRKV 827

Qy        817 STLQNQRSQRSWRLACESASSTEVSGAL-KSTSCSINHPIFPENSEVTFNITFDVDSKAS 875
                ||        ::||    || |  | |:  |::: |||    ||:  : |:    :|
Db        828 EMLQ---PHSRMPVSCEEL--TEGSSLLTKTLKCNVSSPIFKAGQEVSLQVMFNTLLNSS 882

Qy        876 LGNKLLLKANVTSEN-NMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSR 934
              : : |   |  || |             :|| | | ::      || |::||    ::
Db        883 WEDFVELNGTVHCENENSSLQEDNSAATHIPVLYPVNILTKEQENSTLYISFTPKGPKTQ 942

Qy        935 VMQHQYQVSNLGQRSLPISLVFLVPVRLNQTVIWDRPQ----VTFSENLSS----TCHTK 986
              :|| |||     |  | :   :|  |   |   ||     :|:: :: :    |||::
Db        943 QVQHVYQV-----RIQPSAYDHNMPT-LEALVGVPRPHSEDLITYTWSVQTDPLVTCHSE 996

Qy        987 E-RLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTS 1045
              : : ||     ||    | |         : :| | |   : |    : | :     || |
Db        997 DLKRPSSE---AEQPCLPGV--------QFRCPIVF---RWEILIQVTGTVELSKEIKAS 1042

Qy       1046 HNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLL 1105
              : | : |:  : || |    | |   |   :|   ||:        | : | | :|||:||
Db       1043 -STLSLCSSLSVSFNSSKHFHLYGSKA-SEAQVLVKVDLIHEKEMLHVYVLSGIGGLVLL 1100

Qy       1106 ALITAALYKLGFFKRQYKDMM-SEGGPPGAEP 1136
                ||  ||||:|||||  |: | ::|| |   |
Db       1101 FLIFLALYKVGFFKRNLKEKMEADGGVPNGSP 1132


RESULT 6
A53213
integrin alpha-E chain - human
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 09-Jul-2004
C;Accession: A53213
R;Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6016-6025, 1994
A;Title: Molecular cloning of the human mucosal lymphocyte integrin alpha(E) subunit. Un
A;Reference number: A53213; MUID:94164962; PMID:8119947
A;Accession: A53213
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1179 <SHA>
A;Cross-references: UNIPROT:P38570; GB:L25851; NID:g457244; PID:g457245
C;Genetics:
A;Gene: GDB:ITGAE
A;Cross-references: GDB:330801
A;Map position: 17p13
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
F;199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

  Query Match             19.6%;  Score 1149;  DB 2;  Length 1179;
  Best Local Similarity   29.1%;  Pred. No. 1.5e-70;
  Matches  343;  Conservative  211;  Mismatches  451;  Indels  172;  Gaps   39;

Qy         45 GSLYQCDYSTGS--CEPI-RLQVP------VEAVNMSLGLSLAATTSPPQLLACGPTVHQ 95
              | |::|       | |:  : :|       |  |     |: :       |:|   |  |
Db         65 GPLHRCSLVQDEILCHPVEHVPIPKGRHRGVTVVRSHHGVLICI-----QVLVRRP--HS 117

Qy         96 TCSENTYVKGLCFLFGSNLRQQPQ---------------------------------- 119
                || |   | | | |  :|| | |
Db        118 LSSELT---GTCSLLGPDLRPQAQANFFDLENLLDPDARVDTGDCYSNKEGGGEDDVNTA 174

Qy        120 KFPEALRGCPQED-------------SDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQL-- 164
              :   ||    :||              ::|| ::|||||  | ||:| |:|:| :|
Db        175 RQRRALEKEEEEDKEEEEDEEEEEAGTEIAIILDGSGSIDPPDFQRAKDFISNMMRNFYE 234

Qy        165 KKSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLN 224
              |  :  |:|:||    :  |  :: |:     : |: |||:   | ||: :: |:  :
Db        235 KCFECNFALVQYGGVIQTEFDLRDSQDVMASLARVQNITQVGSVTKTASAMQHVLDSIFT 294

Qy        225 ITNGARKNAFKILIVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQEL 284
                :|:|: | |:::|:|||  | |||     ||      :|| |: ||||: |:| :: :||
Db        295 SSHGSRRKASKVMVVLTDGGIFEDPLNLTTVINSPKMQGVERFAIGVGEEFKSARTAREL 354

Qy        285 NTIASKPPRDHVFQVNNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAIT 344
              | ||| |   | |:| |: ||  : ::||  | ::|||       :  ::::| |||| |
Db        355 NLIASDPDETHAFKVTNYMALDGLLSKLRYNIISMEGT---VGDALHYQLAQIGFSAQIL 411

Qy        345 SNGP-LLSTVGSYDWAGGVFLY-TSKEKSTFINMTRVDSDMNDA----YLGYAAAIILRN 398
                 ||  ||::||:||  || |   :  |:| |   :   :|    ||||| |:: :
Db        412 DERQVLLGAVGAFDWSGGALLYDTRSRRGRFLNQTAAAAADAEAAQYSYLGYAVAVLHKT 471

Qy        399 RVQSLVLGAPRYQHIGLVAMFR-QNTGMWESNANV-KGTQIGAYFGASLCSVDVDSNGST 456
              | | |||:|:| |  |:|  |  |   |   | :| |:|:|||: || ||:| :|||
Db        472 CSLSYVAGAPQYKHHG--AVFELQKEGREASFLPVLEGEQMGSYFGSELCPVDIDMDGST 529

Qy        457 DLVLIGAPHYYEQTRGGQVSVCPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVN 516
              | :|: || |:      |:| |  |   |   :  :| |  |   ||| |:  :||::
Db        530 DFLLVAAPFYHVHGEEGRVYVYRLSE-QDGSFSLARILSGHPGFTNARFGFAMAAMGDLS 588

Qy        517 GDKLTDVAIGAP----GEEDNR--GAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQ 570
              |||||||||||     | :|    |:||:::|    |:| | ||||  | ::| |||||
Db        589 QDKLTDVAIGAPLEGFGADDGASFGSVYIYNG-HWDGLSASPSQRIRASTVAPGLQYFGM 647

Qy        571 SLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKG 630
              |::|| |:: ||| |:|||  |  :: ||:||:|:|  | | |  :           :
Db        648 SMAGGFDISGDGLADITVGTLGQAVVFRSRPVVRLKVSMAFTPSALP----------IGF 697

Qy        631 KEAGEVRVCLHVQKSTRDRLREGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVL 690
                  ||:|  :  |       |  :::: : | :| |:   |   :: ::      :
Db        698 NGVVNVRLCFEI-SSVTTASESGLREALLNFTLDVDVGKQRRRLQCSDVRSCLGCLREWS 756

Qy        691 GLTQTCETLKLQLPN----CIEDPVSPIVLRLNFSLVGTPLSAFGNLRPVLAEDAQRLFT 746
              :| || | | :|     | ||  |   ::::: |  ||      : :|:|    :
Db        757 SGSQLCEDL-LLMPTEGELCEEDCFSNASVKVSYQL-QTPEGQTDHPQPILDRYTEPFAI 814

Qy        747 ALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFF 806
              |:|| | |   |  :| :  : :|   ||||  :|  : : : | ||||| | :
Db        815 FQLPYEKACKNKLFCVAELQLA-TTVSQQELVVGLTKELTLNINLTNSGEDSYMTSMALN 873

Qy        807 FPLDLSYRKVSTLQNQRSQRSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNI 866
              :| :|       |:  :   |  : |:          ::   :| | ||:  : |    ::
Db        874 YPRNLQ------LKRMQKPPSPNIQCDDPQPV---ASVLIMNCRIGHPVL-KRSSAHVSV 923

Qy        867 TFDVDSKASLGNKLLLKANVTSENN----MPRTNKTEFQ---LELPVKYAVYMVVTSHGV 919
              : ::  |       :   ||: |       |:  :|:   : :  | ::  | |  |:
Db        924 VWQLEENAFPNRTADITVTVTNSNERRSLANETHTLQFRHGFVAVLSKPSIMYVNTGQGL 983

Qy        920 S--TKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPVRLNQTVIWDRPQVTFSE 977
              |   ::|     ||   :  :||          |  || :|    :     ::| ::
Db        984 SHHKEFLFHVHGEN---LFGAEYQ-----------LQICVPTKLRGLQVAAVKKLTRTQ 1028

Qy        978 NLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLS 1037
              :    ::||  ::|  :  : :   :| ||               :: | |:   :|
Db       1029 ASTVCTWSQERACAYSS-VQHVEEWHSVSCVIA-------------SDKENVTVAAEIS 1073

Qy       1038 FDWYIKTSHNHLLIVST------AEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPL 1091
              :|      |:  |:           || || |::  |  :    |::        |  : |
Db       1074 WD------HSEELLKDVTELQILGEISFNKSLYEGLNAENH--RTKITVVFLKDEKYHSL 1125

Qy       1092 PLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              |:|:  ||||||:| :|   |:| |||||:|: :  |
Db       1126 PIIIKGSVGGLLVLIVILVILFKCGFFKRKYQQLNLE 1162


RESULT 7
A45226
integrin alpha-1 chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A45226
R;Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.
J. Biol. Chem. 268, 2989-2996, 1993
A;Title: Expression of native and truncated forms of the human integrin alpha 1 subunit
```

A;Reference number: A45226; MUID:93155124; PMID:8428973
A;Accession: A45226
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1151 <BRI>
A;Cross-references: UNIPROT:P56199
A;Experimental source: hepatoblastoma cell line HepG2
A;Note: sequence extracted from NCBI backbone (NCBIP:124326)
F;142-317/Domain: von Willebrand factor type A repeat homology <VWA1>

Query Match 18.8%; Score 1102.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 2.4e-67;
Matches 344; Conservative 211; Mismatches 487; Indels 195; Gaps 44;

```
Qy          1 FNLDTENAMTFQENARG-FGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQCDYSTGS 56
              ||:| :|:|||       || :| |    :|  |::|:|       |: | :|:|    |
Db          1 FNVDVKNSMTFSGPVEDMFGYTVQQYENEEGKWVLIGSPLVGQPKNRTGDVYKCPVGRGE 60

Qy         57 CEP-IRLQVPVEA---------VNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGL 106
               | ::| :||            ||: | :|  |      |||||      |    | |:
Db         61 SLPCVKLDLPVNTSIPNVTEVKENMTFGSTL-VTNPNGGFLACGPLYAYRCGHLHYTTGI 119

Qy        107 CFLFGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLK- 165
              |       :     |  :: |  :  ||  ::||| || | |    :  |:: :::::
Db        120 CSDVSPTFQVVNSIAP--VQECSTQ-LDIVIVLDGSNSIYPWD--SVTAFLNDLLKRMDI 174

Qy        166 -KSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLVKPITQLLGR-THTATGVRKVIRELL 223
               :|   ::|| |     |   :: :       | | |  || | || |     :|
Db        175 GPKQTQVGIVQYGENVTHEFNLNKYSSTEEVLVAAKKIVQRGGRQTMTALGTDTARKEAF 234

Qy        224 NITNGARKNAFKILIVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFR-----SE 278
                 |||:   |:::::|||    |     : || : : | :|: | :  ::       :|
Db        235 TEARGARRGVKKVMVIVTDGESH-DNHRLKKVIQDCEDENIQRFSIAILGSYNRGNLSTE 293

Qy        279 KSRQELNTIASKPPRDHVFQVNNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEG 338
              |  :|: :|||:|    | | |::  || ||    | |:|||:| |    |::||| |||| |
Db        294 KFVEEIKSIASEPTEKHFFNVSDELALVTIVKTLGERIFALEATADQSAASFEMEMSQTG 353

Qy        339 FSAAITSNGPLLSTVGSYDWAGGVFLYTSKE-----KSTF-INMTRVDSDMNDAYLGYAA 392
              |||  :  :  :|  ||:||| | | :  :  :       :|| :  |: :  :  :||||
Db        354 FSAHYSQDWVMLGAVGAYDWNGTVVMQKASQIIIPRNTTFNVESTKKNEPL-ASYLGYTV 412

Qy        393 AIILRNRVQSL-VLGAPRYQHIGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSVDVD 451
                :  | : | ||| | | | ::|   |  :    : | |||:|||: | : |:|
Db        413 NSATASSGDVLYIAGQPRYNHTGQVIIYRMEDGNIKILQTLSGEQIGSYFGSILTTTDID 472

Qy        452 SNGSTDLVLIGAPHY---------------YEQTR-GGQVSVCPLPRGQRARWQCDAVLY 495
               : :||::|:||| |                    |||   |:|: |: :  :  | ::
Db        473 KDSNTDILLVGAPMYMGTEKEEQGKVYVYALNQTRFEYQMSLEPIKQTCCSSRQHNSCTT 532

Qy        496 GEQGQPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQ 554
               : :| | ||| |:  : |:| |   |: |||| |:|: ||||::|| ||  |   ::|
Db        533 ENKNEPCGARFGTAIAAVKDLNLDGFNDIVIGAPLEDDHGGAVYIYHG-SGKTIRKEYAQ 591

Qy        555 RIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIMEFNPR 614
              ||         |::|||: |  ||  ||| |:|:|   |   |  |: |  || | | |
Db        592 RIPSGGDGKTLKFFGQSIHGEMDLNGDGLTDVTIGGLGGAALFWSRDVAVVKVTMNFEPN 651

Qy        615 EVARNVFECNDQVVKGKEAG--EVRVCLHVQ-KSTRDRLREGQIQSVVTYDLALDSGRPH 671
               :|      |:   ::|||       ||  |: ||  | : |  :|    | : ||| |
Db        652 KVNIQKKNCH---MEGKETVCINATVCFEVKLKSKEDTIYEADLQ----YRVTLDSLRQI 704

Qy        672 SRAVFNET------KNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNFSLVGT 725
              ||: |: |       :| | |:::     |  :  :   :  :| |    : |:|:|
Db        705 SRSFFSGTQERKVQRNITVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDFNLT-D 755

Qy        726 PLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVGGPRE- 784
              | :      |||  :          || |:|||    |  |||:  :     | |:|    :
Db        756 PENG-----PVLDDSLPNSVHEYIPFAKDCGNKEKCISDLSLHVATTEKDLLIVRSQNDK 810

Qy        785 FNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEVSGAL 844
              |||::||:|  : :| |:     :  :| :  :  :|           :|||
Db        811 FNVSLTVKNTKDSAYNTRTIVHYSPNLVFSGIEAIQKD--------SCESN--------- 853
```

```
Qy        845 KSTSCSINHPIFPPENSEVTFNITFDVDSKASLGN-KLLLKANVTSENNMPRTNKTEFQLE 903
               : :| : :|        |||  | |  ::   : |  : | |   ||      :    :
Db        854 HNITCKVGYPFLRRGEMVTFKILFQFNTSYLMENVTIYLSATSDSEEPPETLSDNVVNIS 913

Qy        904 LPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSN------LGQRS-------L 950
              :|||| | :   |   |  :::  |:|    |:     : |      | ::|       |
Db        914 IPVKYEVGLQFYS-SASEYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSFPMPEL 972

Qy        951 PISLVF----------LVPVRLNQTVIWDRPQVTFSENLSSTCHTKE---------RLPS 991
               :|: |          | |  |:           |||  :   |  |          :: :
Db        973 KLSISFPNMTSNGYPVLYPTGLSS----------SENANCRPHIFEDPFSINSGKKMTT 1021

Qy        992 HSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQE-----------------EFNATLK 1033
               :|    |::  :::|:      | |::    | :                    | |::
Db       1022 STD---HLKRGTILDCNTCKFATITCNLTSSDISQVNVSLILWKPTFIKSYFSSLNLTIR 1078

Qy       1034 GNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPL 1093
              | |         | |  |::|::                    :  :   ::    :|  :||
Db       1079 GEL-------RSENASLVLSSSN----------------QKRELAIQISKDGLPGRVPL 1114

Qy       1094 --IVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
                |: |:  ||||| |:  ||:|:|||||  |  | :
Db       1115 WVILLSAFAGLLLLMLLILALWKIGFFKRPLKKKMEK 1151
```

RESULT 8
I45914
integrin alpha 2 subunit - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C;Accession: I45914
R;Kamata, T.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A;Title: Identification of putative ligand binding sites within the I-domain of integri
A;Reference number: A54402; MUID:94193647; PMID:7511592
A;Accession: I45914
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1170 <KAM>
A;Cross-references: UNIPROT:P53710; GB:L25886; NID:g439695; PIDN:AAB59255.1; PID:g43969
F;161-336/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.5%; Score 1084; DB 2; Length 1170;
Best Local Similarity 27.7%; Pred. No. 4.6e-66;
Matches 336; Conservative 216; Mismatches 495; Indels 168; Gaps 47;

```
Qy          1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQC--DYST 54
               :|:    |  |   ::  || :| |     :|: ::||:|       |: | :|:|  | ||
Db         19 YNVGLPKAKIFSGPSSEQFGYAVQQFINPKGNWLLVGSPWSGFPKNRMGDVYKCPVDLST 78

Qy         55 GSCEPIRLQ-------VPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLC 107
               :|| : ||        |     ||||||:|          | |||   | |    |  |:|
Db         79 TTCEKLNLQTSTSMSNVTEMKTNMSLGLTLTRNVGTGGFLTCGPLWAQQCGSQYYTTGVC 138

Qy        108 FLFGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLK-- 165
                 :  |    |  |:: ||    |:  : | | || | |   :| |:   :: |
Db        139 SDVSPDF-QLRTSFAPAVQTCP-SFIDVVVVCDESNSIYPWD--AVKNFLEKFVQGLDIG 194

Qy        166 KSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLVKPITQLL----GRTHTATGVRKVIRE 221
               :||   |:||:   |: |    |::       ::|  :|         |:|   ::
Db        195 PTKTQMGLIQYANNPRVVFNLNTFKSKD---EMIKATSQTFQYGGDLTNTFKAIQYARDT 251

Qy        222 LLNITNGARKNAFKILIVITDGEKFGDPLGYEDVIPEADREGVIRYVIGV-----GDAFR 276
               :  | |  | |:::|:||||   |    : || : ::: ::|: | |      :|
Db        252 AYSTAAGGRPGATKVMVVVTDGESH-DGSKLKAVIDQCNKDNILRFGIAVLGYLNRNALD 310

Qy        277 SEKSRQELNTIASKPPRDHVFQVNNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQ 336
               ::   :|:  ||| |   | | |::     |       : |:||:|||| |   :|: ||||
Db        311 TKNLIKEIKAIASIPTERHFFNVSDEADLLEKAGTIGEQIFSIEGTVQG-GDNFQMEMSQ 369

Qy        337 EGFSAAIT--SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYA 391
              ||||  :  :|  :|  ||:|||:| |   |       |      ::  | |  :||||:
```

```
Db        370 VGFSAEYSPQNNILMLGAVGAYDWSGTVVQKTPHGHLIFSKQAFEQILQDRNHSSYLGYS 429

Qy        392 AAIILRNRVQSLVLGAPRYQHIGLVAMFRQNTGMWESNANV------KGTQIGAYFGASL 445

Db        430 VASISTGNSVHFVAGAPRANYTGQIVLYSVN-----ENGNVTVIQSQRGDQIGSYFGSVL 484

Qy        446 CSVDVDSNGSTDLVLIGAPHYYEQTR--GGQVSVCPLPRGQRARWQCDAVLYGEQGQPWG 503

Db        485 CAVDVNKDTITDVLLVGAPMYMNDLKKEEGRVYLFTITKG-ILNWH--QFLEGPNGLENA 541

Qy        504 RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGS--KL 561

Db        542 RFGSAIAALSDINMDGFNDVIVGSPLENQNSGAVYIYNGHEGM-IRLRYSQKILGSDRAF 600

Qy        562 SPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVF 621

Db        601 SSHLQYFGRSLDGYGDLNGDSITDVSVGAFGQVVQLWSQSIADVSVDASFTPKKI--TLL 658

Qy        622 ECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVTYDLALD----SGRPHSRAVFN 677

Db        659 NKNAEI-------KLKLCF----SAKFRPTNQNNQVAIVYNITIDEDQFSSRVISRGLFK 707

Qy        678 ETKNSTRRQTQVLGLTQTCE--TLKLQLPNCIEDPVSPIVLRLNFSL--VGTPLSAFGNL 733

Db        708 ENNERCLQKTMIVSQAQRCSEYIIHIQEPS---DIISPLNLCMNISLENPGT-------- 756

Qy        734 RPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITF----SFMSLDCLVVGGPREFNVTV 789

Db        757 NPALEAYSETVKVFSIPFHKDCGDDGVCISDLVLNVQQLPATQQQPFIVSNQNKRLTFSV 816

Qy        790 TVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASST-EVSGALKSTS 848

Db        817 QLKNKKESAYNTEIVVDFSENLFF-----------ASWSMPVDGTEVTCQIASSQKSVT 864

Qy        849 CSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEFQLELPVKY 908

Db        865 CNVGYPALKSKQQVTFTINFDFNLQ-NLQNQASISFRALSESQEENMADNSVNLKLSLLY 923

Qy        909 AVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQR----------SLPISLVFLV 958

Db        924 DAEIHIT-RSTNINFYEVSLDGNVSSVV-HSFE--DIGPKFIFSIKVTTGSVPVSMA--- 976

Qy        959 PVRLNQTVIWDRPQVTFSEN---LSSTCHTKE---------------RLPSHSDFLAE- 998

Db        977 ------SVIIHIPQYTKDKNPLMYLTGVHTDQAGDISCEAEINPLKIGQTSSSVSFKSEN 1030

Qy        999 LRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEI- 1057

Db       1031 FRHIKELNCRTASCSNIMCWLRDLQVKGEYFLNVSTRIWNGTFAASTFQTVQLTAAAEID 1090

Qy       1058 LFNDSVFTL------LPGQGAFVRSQTETKVEPPE-VPNPLPLIVGSSVGGLLLLALITA 1110

Db       1091 TYNPQIYVIEENTVTIP----------LTIMKPHEKVEVPTGVIVGSVIAGILLLLALVA 1140

Qy       1111 ALYKLGFFKRQYKDM 1125

Db       1141 ILWKLGFFKRKYEKM 1155
```

RESULT 9
S44142
VLA-2 protein homolog - mouse
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S44142
R;Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze
submitted to the EMBL Data Library, January 1994
A;Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not
A;Reference number: S44142
A;Accession: S44142
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1178 <EDE>

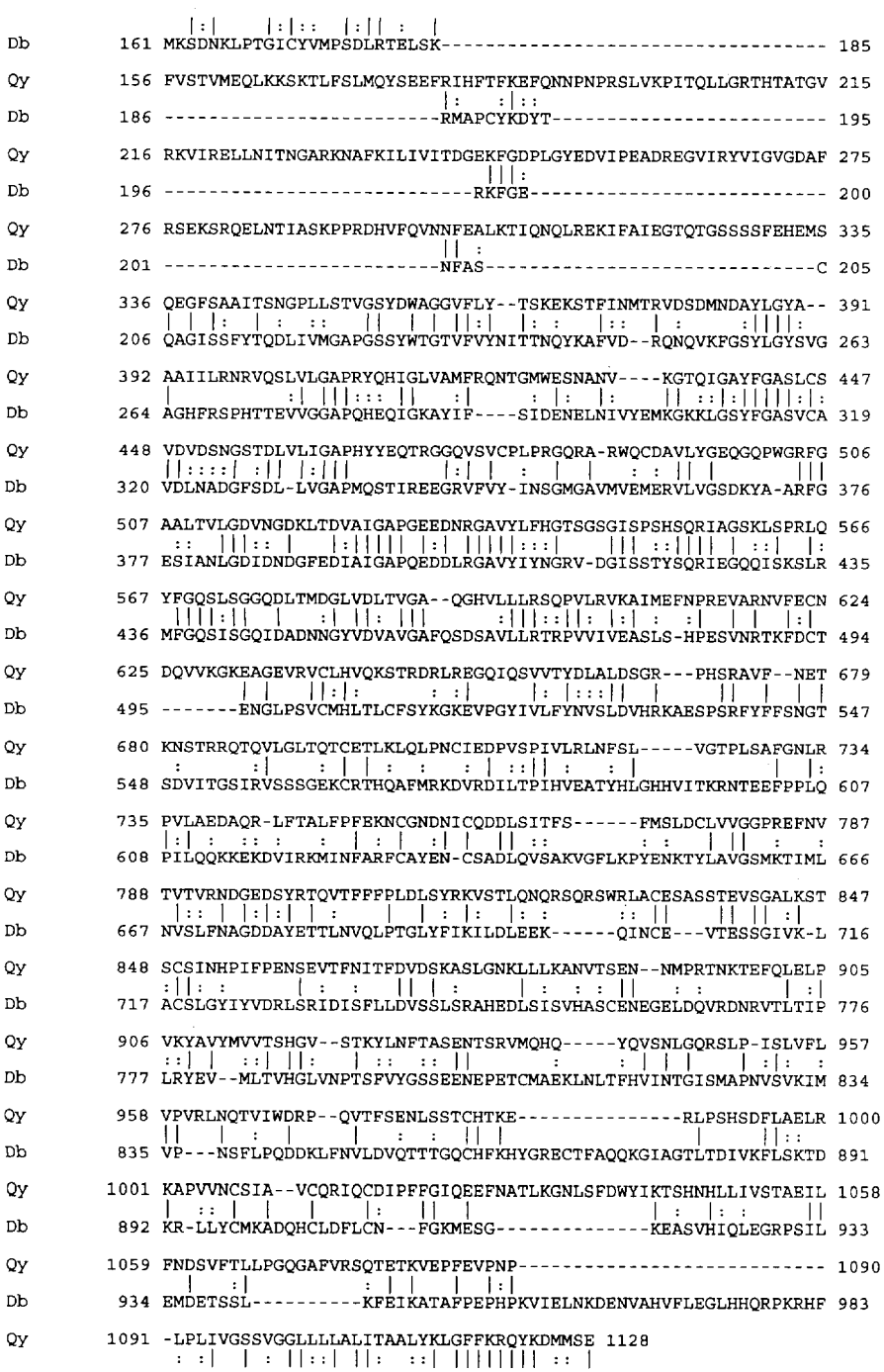
A;Cross-references: UNIPROT:Q62469; EMBL:Z29987; NID:g473098; PIDN:CAA82877.1; PID:g473
F;169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.2%; Score 1072; DB 2; Length 1178;
Best Local Similarity 28.1%; Pred. No. 3.1e-65;
Matches 344; Conservative 206; Mismatches 487; Indels 188; Gaps 44;

```
Qy          1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQC--DYST 54

Db         27 YNVGLPGAKIFSGPSSEQFGYSVQQLTNPQGNWLLVGSPWSGFPENRMGDVYKCPVDLPT 86

Qy         55 GSCEPIRLQ-------VPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLC 107

Db         87 ATCEKLNLQNSASISNVTEIKTNMSLGLTLTRNPGTGGFLTCGPLWAHQCGNQYYATGIC 146

Qy        108 FLFGSNLRQQPQ---KFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQL 164

Db        147 ----SDVSPDFQFLTSFSPAVQACPSL-VDVVVVCDESNSIYP--WEAVKNFLVKFVTGL 199

Qy        165 K--KSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLVKPITQLLG-RTHTATGVRKVIRE 221

Db        200 DIGPKKTQVALIQYANEPRIIFNLNDFETKEDMVQATSETRQHGGDLTNTFRAIEFARDY 259

Qy        222 LLNITNGARKNAFKILIVITDGEKFGDPLGYEDVIPEADREGVIRYVIGV-----GDAFR 276

Db        260 AYSQTSGGRPGATKVMVVVTDGESH-DGSKLKTVIQQCNDDEILRFGIAVLGYLNRNALD 318

Qy        277 SEKSRQELNTIASKPPRDHVFQVNNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQ 336

Db        319 TKNLIKEIKAIASTPTERYFFNVADEAALLEKAGTLGEQIFSIEGTVQG-GDNFQMEMAQ 377

Qy        337 EGFSA--AITSNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYA 391

Db        378 VGFSADYAPQNDILMLGAVGAFDWSGTLVQETSHKPVIFPKQAFDQVLQDRNHSSFLGYS 437

Qy        392 AAIILRNRVQSLVLGAPRYQHIGLVAMFRQNTGMWESNANV----KGTQIGAYFGASLCS 447

Db        438 VAAISTEDGVHFVAGAPRANYTGQIVLYSVNK---QGNVTVIQSHRGDQIGSYFGSVLCS 494

Qy        448 VDVDSNGSTDLVLIGAPHYYEQTR--GGQVSVCPLPRGQRARWQCDAVLYGEQGQPWGRF 505

Db        495 VDVDKDTITDVLLVGAPTYMNDLKKEEGKVYLFTITKGILNQHQ---FLEGPEGTGNARF 551

Qy        506 GAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPR- 564

Db        552 GSAIAALSDINMDGFNDVIVGSPVENENSGAVYIYNGHQGT-IRTKYSQKILGSNGAFRR 610

Qy        565 -LQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFEC 623

Db        611 HLQFFGRSLDGYGDLNGDSITDVSIGALGQVIQLWSQSIADVAIEALFTP---------- 660

Qy        624 NDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVTYDLALD----SGRPHSRAVFNET 679

Db        661 -DKITLLNKDAKITLKLCFRAEFRPAGQNNQV--AILFNMTLDADGHSSRVTSRGVFREN 717

Qy        680 KNSTRRQTQVLGLTQTCET--LKLQLPNCIEDPVSPIVLRLNFSLVGTPLSAFGNLRPVL 737

Db        718 SERFLQKNMVVNEVQKCSEHHISIQKPS---DVVNPLDLRVDISLENPGTS------PAL 768

Qy        738 AEDAQRLFTALFPFEKNCGNDNICQDDLSI---------TFSFMSLDCLVVGGPREFNVT 788

Db        769 EAYSETVKVFSIPFYKECGSDGICISDLILDVQQLPAIQTQSF-----IVSNQNKRLTFS 823

Qy        789 VTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASST-EVSGALKST 847

Db        824 VILKNRGESAYNTVVLAEFSENLFF-----------ASFSMPVDGTEVTCEVGSSQKSV 871

Qy        848 SCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTE--FQLELP 905

Db        872 TCDVGYPALKSEQQVTFTINFDFNLQ-NLQNQAAINFQAFSESQ--ETNKADNSVSLTIP 928

Qy        906 VKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQR----------SLPISLV 955
```

```
Db        929 LLYDAELHLT-RSTNINFYEISSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVSMA 984

Qy        956 FLV---------------------------------PVRLNQTVIWDRPQVTF-SENLS 980
                :                                 |::|  |     | |:| :||
Db        985 LVTIHIPQYTKEKNPLLYLTGIQTDQAGDISCTAEINPLKLPHTA----PSVSFKNENFR 1040

Qy        981 STCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDW 1040
                 ||||                  ::|    |  | | :    :: |:   :   :
Db       1041 ---HTKE----------------LDCRTTSCSNITCWLKDLHMKAEYFINVTTRVWNRT 1080

Qy       1041 YIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVG 1100
              :  ::   : : : |||  ::    ::      :         |  |||    :|:|| :
Db       1081 FAASTFQTVQLTAAAEIDTHNPQLFVIEENAVTIPLMIMKPTEKAEVPT--GVIIGSIIA 1138

Qy       1101 GLLLLALITAALYKLGFFKRQYKDM 1125
              |:|||  :|| |:|||||||||| |
Db       1139 GILLLLAMTAGLWKLGFFKRQYKKM 1163


RESULT 10
A33998
integrin alpha-2 chain precursor - human
N;Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 09-Jul-2004
C;Accession: A33998; B56793; A53117
R;Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet
A;Reference number: A33998; MUID:89308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: UNIPROT:P17301; GB:X17033; NID:g33906; PIDN:CAA34894.1; PID:g33907
A;Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R;Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIc, GPIc*, GPIIa an
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16,'V',18-21 <ZUT>
A;Cross-references: GB:L24121; NID:g400342; PIDN:AAA16619.2; PID:g4583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TMM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

  Query Match           18.2%;  Score 1069;  DB 2;  Length 1181;
  Best Local Similarity  27.0%;  Pred. No. 5e-65;
  Matches  331;  Conservative  212;  Mismatches  494;  Indels  188;  Gaps   43;

Qy          1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQC--DYST 54
              :|:    |  |   ::  || :| |    :|: ::||:|      |: | :|:|  | ||
Db         30 YNVGLPEAKIFSGPSSEQFGYAVQQFINPKGNWLLVGSPWSGFPENRMGDVYKCPVDLST 89

Qy         55 GSCEPIRLQ-------VPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLC 107
                                                                  :|| : ||       |     ||||||  |          | |||    | |     |  |:|
Db         90 ATCEKLNLQTSTSIPNVTEMKTNMSLGLILTRNMGTGGFLTCGPLWAQQCGNQYYTTGVC 149

Qy        108 FLFGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLK-- 165
                :  |    |  | : ||     |: : | || | | |   :| |:   :: |
Db        150 SDISPDF-QLSASFSPATQPCPSL-IDVVVVCDESNSIYPWD--AVKNFLEKFVQGLDIG 205

Qy        166 KSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLVKPITQLLG-RTHTATGVRKVIRELLN 224
              :||   |:||:   |: |    ::          :|  |  |:|   ::   :   :
Db        206 PTKTQVGLIQYANNPRVVFNLNTYKTKEEMIVATSQTSQYGGDLTNTFGAIQYARKYAYS 265

Qy        225 ITNGARKNAFKILIVITDGEKFGDPLGYEDVIPEADREGVIRYVIGV-----GDAFRSEK 279
              :| |::| |:::|:|||| |    : || : : : ::|: | |       :|  ::
Db        266 AASGGRRSATKVMVVVTDGESH-DGSMLKAVIDQCNHDNILRFGIAVLGYLNRNALDTKN 324

Qy        280 SRQELNTIASKPPRDHVFQVNNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGF 339
              :|:  ||| |   : | |::  ||       | |:||:|||| |   :|: |||| ||
Db        325 LIKEIKAIASIPTERYFFNVSDEAALLEKAGTLGEQIFSIEGTVQG-GDNFQMEMSQVGF 383

Qy        340 SAAITSNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYAAAI 394
              ||  :|    :|  ||:: |:| :  ||     |      ::  | |  :|||| |
Db        384 SADYSSQNDILMLGAVGAFGWSGTIVQKTSHGHLIFPKQAFDQILQDRNHSSYLGYSVAA 443

Qy        395 ILRNRVQSLVLGAPRYQHIGLVAMFRQNTGMWESNANV------KGTQIGAYFGASLCSV 448
              |       | ||||  : | : ::  |       | |:       :| |||:|||: ||||
Db        444 ISTGESTHFVAGAPRANYTGQIVLYSVN-----ENGNITVIQAHRGDQIGSYFGSVLCSV 498

Qy        449 DVDSNGSTDLVLIGAPHYYEQTR--GGQVSVCPLPRGQRARWQCDAVLYGEQGQPWGRFG 506
              ||| :  ||::|:||| |    :   |:| : :|   : |    | | :|    |||
Db        499 DVDKDTITDVLLVGAPMYMSDLKKEEGRVYLFTIKKGILGQHQ---FLEGPEGIENTRFG 555

Qy        507 AALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGS--KLSPR 564
              :|:  | |:| |   || :|:|  |  | ||||:::|  |: |   :||:| ||
Db        556 SAIAALSDINMDGFNDVIVGSPLENQNSGAVYIYNGHQGT-IRTKYSQKILGSDGAFRSH 614

Qy        565 LQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECN 624
              |||||:|| |  ||  | : |:::|| | |: | || :  |      | | ::   :   |
Db        615 LQYFGRSLDGYGDLNGDSITDVSIGAFGQVVQLWSQSIADVAIEASFTPEKI--TLVNKN 672

Qy        625 DQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVTYDLALD----SGRPHSRAVFNETK 680
              |::      :::|     | :  |  :   |  : |:: ||     | |  || :| |
Db        673 AQII-------LKLCF----SAKFRPTKQNNQVAIVYNITLDADGFSSRVTSRGLFKENN 721

Qy        681 NSTRRQTQVLGLTQTC--ETLKLQLPNCIEDPVSPIVLRLNFSLVGTPLSAFGNLRPVLA 738
                ::  |:  |:|    : :| |:   | |: : ||:: ||      |      | |
Db        722 ERCLQKNMVVNQAQSCPEHIIYIQEPS---DVVNSLDLRVDISLENPGTS------PALE 772

Qy        739 EDAQRLFTALFPFEKNCGNDNICQDDLSITF----SFMSLDCLVVGGPREFNVTVTVRND 794
              ::       || |:|| | :|  || :      :      :|    :    :||::|
Db        773 AYSETAKVFSIPFHKDCGEDGLCISDLVLDVRQIPAAQEQPFIVSNQNKRLTFSVTLKNK 832

Qy        795 GEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASST-EVSGALKSTSCSINH 853
              | :| | :   |  :| :                |: |  :    | :|: : || :| : :
Db        833 RESAYNTGIVVDFSENLFF-----------ASFSLPVDGTEVTCQVAASQKSVACDVGY 880

Qy        854 PIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEFQLELPVKYAVYMV 913
              |      :||| | || : : :| |:  |    ||:                |::|: |   :
Db        881 PALKREQQVTFTINFDFNLQ-NLQNQASLSFQALSESQEENKADNLVNLKIPLLYDAEI- 938

Qy        914 VTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQR----------SLPISLVFLV----- 958
                |   : :||    :   |    :   ::| :           |:|:|:  ::
Db        939 ---HLTRSTNINFYEISSDGNVPSIVHSFEDVGPKFIFSLKVTTGSVPVSMATVIIHIPQ 995

Qy        959 ----------------------------PVRLNQTVIWDRPQVTF-SENLSSTCHTKER 988
                                         |::: ||       |:| |||     ||||
Db        996 YTKEKNPLMYLTGVQTDKAGDISCNADINPLKIGQT----SSSVSFKSENFR---HTKE- 1047

Qy        989 LPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNH 1048
                                  :||  | |  : |  :    :: |:   :   :   ::
Db       1048 ----------------LNCRTASCSNVTCWLKDVHMKGEYFVNVTTRIWNGTFASSTFQT 1091

Qy       1049 LLIVSTAEI-LFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNP-----LP--LIVGSSVG 1100
              : : : |||  :|  :: :          :  |   |  :  |     :|  :|:|| :
```

```
Db       1092 VQLTAAAEINTYNPEIYVI----------EDNTVTIPLMIMKPDEKAEVPTGVIIGSIIA 1141

Qy       1101 GLLLLALITAALYKLGFFKRQYKDM 1125
              |:|||  :  | |:|||||||:|: |
Db       1142 GILLLLALVAILWKLGFFKRKYEKM 1166


RESULT 11
A35854
integrin alpha-1 chain precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 09-Jul-2004
C;Accession: A35854; S11243
R;Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto,
J. Cell Biol. 111, 709-720, 1990
A;Title: Molecular cloning of the rat integrin alpha-1-subunit: a receptor for laminin a
A;Reference number: A35854; MUID:90338125; PMID:2380249
A;Accession: A35854
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1180 <IGN>
A;Cross-references: UNIPROT:P18614; GB:X52140; NID:g56493; PIDN:CAA36384.1; PID:g56494
C;Keywords: cell adhesion; cytoskeleton; transmembrane protein
F;170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

  Query Match             18.0%;  Score 1060;  DB 2;  Length 1180;
  Best Local Similarity   27.4%;  Pred. No. 2.1e-64;
  Matches  344;  Conservative  201;  Mismatches  478;  Indels  234;  Gaps   47;

Qy          1 FNLDTENAMTFQENARG-FGQSVVQL---QGSRVVVGAPQEIVAANQRGSLYQCDYSTGS 56
              ||:| :|:|:|       || :| |    :|  |::|:|       : | :|:|
Db         29 FNVDVKNSMSFSGPVEDMFGYTVQQYENEEGKWVLIGSPLVGQPKARTGDVYKCPVGRER 88

Qy         57 CEP-IRLQVPVEA---------VNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGL 106
                | ::| :||            ||: | :|  |      |||||      |    |  |:
Db         89 AMPCVKLDLPVNTSIPNVTEIKENMTFGSTL-VTNPNGGFLACGPLYAYRCGHLHYTTGI 147

Qy        107 CFLFGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLK- 165
              |       :      |  :: |  :  ||  ::||| || |  :  :  |:: :::::
Db        148 CSDVSPTFQVVNSFAP--VQECSTQ-LDIVIVLDGSNSIYP--WESVIAFLNDLLKRMDI 202

Qy        166 -KSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLVKPITQLLG-RTHTATGVRKVIRELL 223
               :|   ::|| |      |   :: :          | :  | :| || |:    :|
Db        203 GPKQTQVGIVQYGENVTHEFNLNKYSSTEEVLVAANKIGRQGGLQTMTALGIDTARKEAF 262

Qy        224 NITNGARKNAFKILIVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFR-----SE 278
                 |||:   |:::::|||    |    : || : : | : |: | :   :       :|
Db        263 TEARGARRGVKKVMVIVTDGESH-DNYRLKQVIQDCEDENIQRFSIAILGHYNRGNLSTE 321

Qy        279 KSRQELNTIASKPPRDHVFQVNNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEG 338
              |  :|: :|||:|   | | |::  || ||   | |:|||:| |    |::||| |||| |
Db        322 KFVEEIKSIASEPTEKHFFNVSDELALVTIVKALGERIFALEATADQSAASFEMEMSQTG 381

Qy        339 FSAAITSNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMND---AYLGYAAA 393
              |||  : :  :|  ||:||| | | :  : :      | |      : ||:   :||||
Db        382 FSAHYSQDWVMLGAVGAYDWNGTVVMQKANQMVIPHNTTFQTEPAKMNEPLASYLGYTVN 441

Qy        394 IILRNRVQSLVLGAPRYQHIGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSVDVDSN 453
                    : | ||| | | | :::   |        : | |||:|||: | ::|:| :
Db        442 SATIPGDVLYIAGQPRYNHTGQVVIYKMEDGNINILQTLGGEQIGSYFGSVLTTIDIDKD 501

Qy        454 GSTDLVLIGAPHY---------------YEQTR-GGQVSVCPLPRGQRARWQCDAVLYGE 497
              |||:|:||| |                  |||   |:|: |: :   :  : ::
Db        502 SYTDLLLVGAPMYMGTEKEEQGKVYVYAVNQTRFEYQMSLEPIRQTCCSSLKDNSCTKEN 561

Qy        498 QGQPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQRI 556
              : :| | ||| |:  : |:| |    || |||| |:|: ||||::|| ||  |  :::|||
Db        562 KNEPCGARFGTAIAAVKDLNVDGFNDVVIGAPLEDDHAGAVYIYHG-SGKTIREAYAQRI 620

Qy        557 AGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQPVLRVKAIMEFNPREV 616
                    |::|||: |  ||  ||| |:|:|  |    |  :: |  ||  | | | :|
Db        621 PSGGDGKTLKFFGQSIHGEMDLNGDGLTDVTIGGLGGAALFWARDVAVVKVTMNFEPNKV 680

Qy        617 ARNVFECNDQVVKGKEAG--EVRVCLHVQ-KSTRDRLREGQIQSVVTYDLALDSGRPHSR 673
                      |    |:|||       :| ||: ||  | : |  :|    | : ||| |  ||
Db        681 NIQKKNCR---VEGKETVCINATMCFHVKLKSKEDSIYEADLQ----YRVTLDSLRQISR 733

Qy        674 AVFNET------KNSTRRQTQVLGLTQTCETLKLQLPNCI---------EDPVSPIVLRL 718
              : |: |      :| | |:::                |||         |   : : |
Db        734 SFFSGTQERKIQRNITVRESE-----------------CIRHSFYMLDKHDFQDSVRVTL 776

Qy        719 NFSLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLV 778
              :|:|   | :      |||  :          || |:|||   |  ||::  |      |:
Db        777 DFNLT-DPENG-----PVLDDALPNSVHEHIPFAKDCGNKERCISDLTLNVSTTEKSLLI 830

Qy        779 VGGPRE-FNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASS 837
              |    : |||::||:| |: :| |:       :| :  :  :|          :|||
Db        831 VKSQHDKFNVSLTVKNKGDSAYNTRTVVQHSPNLIFSGIEEIQKD--------SCESN-- 880

Qy        838 TEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLL-LKANVTSENNMPRTN 896
                    :: :| : :|          ||| | |  ::     | :: | |   ||  :   |
Db        881 -------QNITCRVGYPFLRAGETVTFKIIFQFNTSHLSENAIIHLSATSDSEEPLESLN 933

Qy        897 KTEFQLELPVKYAV----YMVVTSHGVST-------KYLNFTASENTSRVMQHQYQVSNL 945
                |  : :|||| |    |   : | :|        :::| |  |:    :   | :
Db        934 DNEVNISIPVKYEVGLQFYSSASEHHISVAANETIPEFINST--EDIGNEINVFYTIRKR 991

Qy        946 GQRSLP---ISLVF----------LVPVRLNQTVIWD-------RP------------Q 972
              |   :|   :|: |          | |:       |        ||             :
Db        992 GHFPMPELQLSISFPNLTADGYPVLYPIG------WSSSDNVNCRPRSLEDPFGINSGKK 1045

Qy        973 VTFS----------ENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI--- 1019
              :| |           :: ||||                 : ||:      | ::
Db       1046 MTISKSEVLKRGTIQDCSSTC-----------------GVATITCSLLPSDLSQVNVSLL 1088

Qy       1020 ---PFFGIQEEF---NATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAF 1073
                  | | |:  |   | ||:| |        | |  | :|::
Db       1089 LWKPTF-IRAHFSSLNLTLRGELK-------SENSSLTLSSSN----------------- 1123

Qy       1074 VRSQTETKVEPFEVPNPLPL--IVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
                : :   ::    :|  :|| |: |:  ||||| |:  ||:|:|||||  |  | :
Db       1124 RKRELAIQISKDGLPGRVPLWVILLSAFAGLLLLMLLILALWKIGFFKRPLKKKMEK 1180


RESULT 12
A41131
lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
N;Alternate names: integrin alpha-4
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004
C;Accession: A41131; S16742
R;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.L.
J. Cell Biol. 115, 1149-1158, 1991
A;Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte
A;Reference number: A41131; MUID:92064645; PMID:1840602
A;Accession: A41131
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1039 <NEU>
A;Cross-references: UNIPROT:Q00651; EMBL:X53176; NID:g51484; PIDN:CAA37316.1; PID:g5148
C;Superfamily: integrin alpha-4 chain
C;Keywords: cytoskeleton; transmembrane protein

  Query Match             11.3%;  Score 663;  DB 2;  Length 1039;
  Best Local Similarity   22.7%;  Pred. No. 4e-37;
  Matches  281;  Conservative  187;  Mismatches  403;  Indels  368;  Gaps   48;

Qy          1 FNLDTENAMTFQ-ENARGFGQSVV-QLQGSR--VVVGAPQEIVAAN----QRGSLYQCDY 52
              :||| |||: :|  :    || |||     ||:  ::|||       :|       |::|:|
Db         41 YNLDPENALLYQGPSGTLFGYSVVLHSHGSKRWLIVGAPTASWLSNASVVNPGAIYRCGI 100

Qy         53 STG---SCEPIRLQVP----------VEAVNMSLGLSLAATTSP-PQLLACG---PTVHQ 95
                 :|| ::   |            |  |  ||::|:          ::  ||     :
Db        101 RKNPNQTCEQLQSGSPSGEPCGKTCLEERDNQWLGVTLSRQPGENGSIVTCGHRWKNIFY 160

Qy         96 TCSENTYVKGLCFLFGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKE 155
```



Db 161 MKSDNKLPTGICYVMPSDLRTELSK---------- 185

Qy 156 FVSTVMEQLKKSKTLFSLMQYSEEFRIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGV 215

Db 186 ----------RMAPCYKDYT---------- 195

Qy 216 RKVIRELLNITNGARKNAFKILIVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAF 275

Db 196 ----------RKFGE---------- 200

Qy 276 RSEKSRQELNTIASKPPRDHVFQVNNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMS 335

Db 201 ----------NFAS----------C 205

Qy 336 QEGFSAAITSNGPLLSTVGSYDWAGGVFLY--TSKEKSTFINMTRVDSDMNDAYLGYA-- 391

Db 206 QAGISSFYTQDLIVMGAPGSSYWTGTVFVYNITTNQYKAFVD--RQNQVKFGSYLGYSVG 263

Qy 392 AAIILRNRVQSLVLGAPRYQHIGLVAMFRQNTGMWESNANV----KGTQIGAYFGASLCS 447

Db 264 AGHFRSPHTTEVVGGAPQHEQIGKAYIF----SIDENELNIVYEMKGKKLGSYFGASVCA 319

Qy 448 VDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPLPRGQRA-RWQCDAVLYGEQGQPWGRFG 506

Db 320 VDLNADGFSDL-LVGAPMQSTIREEGRVFVY-INSGMGAVMVEMERVLVGSDKYA-ARFG 376

Qy 507 AALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQ 566

Db 377 ESIANLGDIDNDGFEDIAIGAPQEDDLRGAVYIYNGRV-DGISSTYSQRIEGQQISKSLR 435

Qy 567 YFGQSLSGGQDLTMDGLVDLTVGA--QGHVLLLRSQPVLRVKAIMEFNPREVARNVFECN 624

Db 436 MFGQSISGQIDADNNGYVDVAVGAFQSDSAVLLRTRPVVIVEASLS-HPESVNRTKFDCT 494

Qy 625 DQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVTYDLALDSGR---PHSRAVF--NET 679

Db 495 -------ENGLPSVCMHLTLCFSYKGKEVPGYIVLFYNVSLDVHRKAESPSRFYFFSNGT 547

Qy 680 KNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNFSL-----VGTPLSAFGNLR 734

Db 548 SDVITGSIRVSSSGEKCRTHQAFMRKDVRDILTPIHVEATYHLGHHVITKRNTEEFPPLQ 607

Qy 735 PVLAEDAQR-LFTALFPFEKNCGNDNICQDDLSITFS------FMSLDCLVVGGPREFNV 787

Db 608 PILQQKKEKDVIRKMINFARFCAYEN-CSADLQVSAKVGFLKPYENKTYLAVGSMKTIML 666

Qy 788 TVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEVSGALKST 847

Db 667 NVSLFNAGDDAYETTLNVQLPTGLYFIKILDLEEK------QINCE---VTESSGIVK-L 716

Qy 848 SCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSEN--NMPRTNKTEFQLELP 905

Db 717 ACSLGYIYVDRLSRIDISFLLDVSSLSRAHEDLSISVHASCENEGELDQVRDNRVTLTIP 776

Qy 906 VKYAVYMVVTSHGV--STKYLNFTASENTSRVMQHQ-----YQVSNLGQRSLP-ISLVFL 957

Db 777 LRYEV--MLTVHGLVNPTSFVYGSSEENEPETCMAEKLNLTFHVINTGISMAPNVSVKIM 834

Qy 958 VPVRLNQTVIWDRP--QVTFSENLSSTCHTKE----------RLPSHSDFLAELR 1000

Db 835 VP---NSFLPQDDKLFNVLDVQTTTGQCHFKHYGRECTFAQQKGIAGTLTDIVKFLSKTD 891

Qy 1001 KAPVVNCSIA--VCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEIL 1058

Db 892 KR-LLYCMKADQHCLDFLCN---FGKMESG----------KEASVHIQLEGRPSIL 933

Qy 1059 FNDSVFTLLPGQGAFVRSQTETKVEPFEVPNP---------- 1090

Db 934 EMDETSSL----------KFEIKATAFPEPHPKVIELNKDENVAHVFLEGLHHQRPKRHF 983

Qy 1091 -LPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128

Db 984 TIIIITISLLLGLIVLLLISCVMWKAGFFKRQYKSILQE 1022

RESULT 13
S06046
integrin alpha-4 chain precursor - human
N;Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain
C;Species: Homo sapiens (man)
C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C;Accession: S06046; A39355; D28018
R;Takada, Y.; Elices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989
A;Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other inte
A;Reference number: S06046; MUID:89356603; PMID:2788572
A;Accession: S06046
A;Molecule type: mRNA
A;Residues: 1-1038 <TAK>
A;Cross-references: UNIPROT:P13612; GB:X16983; EMBL:X15356; NID:g33945; PIDN:CAA34852.1
R;Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A;Title: Characterization of the alpha-4 integrin gene promoter.
A;Reference number: A39355; MUID:91239513; PMID:2034655
A;Accession: A39355
A;Molecule type: DNA
A;Residues: 1-93 <ROS>
A;Cross-references: GB:M62841
R;Takada, Y.; Strominger, J.L.; Hemler, M.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987
A;Title: The very late antigen family of heterodimers is part of a superfamily of molecu
A;Reference number: A94151; MUID:87204112; PMID:3033641
A;Accession: D28018
A;Molecule type: protein
A;Residues: 40-50,'E',52-53 <TA2>
C;Genetics:
A;Gene: GDB:ITGA4; CD49D
A;Cross-references: GDB:128032; OMIM:192975
A;Map position: 2q31-2q32
C;Superfamily: integrin alpha-4 chain
C;Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transme
F;1-39/Domain: signal sequence #status predicted <SIG>
F;40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.8%; Score 633; DB 2; Length 1038;
Best Local Similarity 25.2%; Pred. No. 4.7e-35;
Matches 248; Conservative 155; Mismatches 370; Indels 212; Gaps 39;

Qy 272 GDAFRSEKSRQELNTIASKPPRD----------HVFQVNNFEALKT---------IQN 310

Db 121 GKTCLEERDNQWLGVTLSRQPGENGSIVTCGHRWKNIFYIKNENKLPTGGCYGVPPDLRT 180

Qy 311 QLREKI----------FAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSY 356

Db 181 ELSKRIAPCYQDYVKKFGENFA----------SCQAGISSFYTKDLIVMGAPGSS 225

Qy 357 DWAGGVFLY---TSKEKSTFINMTRVDSDMNDAYLGYA--AAIILRNRVQSLVLGAPRYQ 411

Db 226 YWTGSLFVYNITTNKYKAFLDKQNQVKF---GSYLGYSVGAGHFRSQHTTEVVGGAPQHE 282

Qy 412 HIGLVAMFRQNTGMWESNANV----KGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYY 467

Db 283 QIGKAYIF----SIDEKELNILHEMKGKKLGSYFGASVCAVDLNADGFSDL-LVGAPMQS 337

Qy 468 EQTRGGQVSVCPLPRGQRARWQC-DAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIG 526

Db 338 TIREEGRVFVY-INSGSGAVMNAMETNLVGSDKYA-ARFGESIVNLGDIDNDGFEDVAIG 395

Qy 527 APGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDL 586

Db 396 APQEDDLQGAIYIYNGRA-DGISSTFSQRIEGLQISKSLSMFGQSISGQIDADNNGYVDV 454

Qy 587 TVGA--QGHVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQK 644

Db 455 AVGAFRSDSAVLLRTRPVVIVDASLS-HPESVNRTKFDC-------VENGWPSVCIDLTL 506

```
Qy        645 STRDRLREGQIQSVVTYDLALDSGR-----PHSRAVFNETKNSTRRQTQVLGLTQTCETL 699
Db        507 CFSYKGKEVPGYIVLFYNMSLDVNRKAESPPRFYFSSNGTSDVITGSIQVSSREANCRTH 566

Qy        700 KLQLPNCIEDPVSPIVLRLNFSLVGTPLS-----AFGNLRPVLAEDAQR-LFTALFPFEK 753
Db        567 QAFMRKDVRDILTPIQIEAAYHLGPHVISKRSTEEFPPLQPILQQKKEKDIMKKTINFAR 626

Qy        754 NCGNDNICQDDLSIT--FSFM----SLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFF 807
Db        627 FCAHEN-CSADLQVSAKIGFLKPHENKTYLAVGSMKTLMLNVSLFNAGDDAYETTLHVKL 685

Qy        808 PLDLSYRKVSTLQNQRSQRSWRLACESASSTEVSGALKSTSCSINHPIFPPENSEVTFNIT 867
Db        686 PVGLYFIKILELEEK------QINCE---VTDNSGVVQ-LDCSIGYIYVDHLSRIDISFL 735

Qy        868 FDVDSKASLGNKLLLKANVTSEN--NMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLN 925
Db        736 LDVSSLSRAEEDLSITVHATCENEEEMDNLKHSRVTVAIPLKYEVKLTVHGFVNPTSFVY 795

Qy        926 FTASENTSRV-----MQHQYQVSNLGQRSLP-ISLVFLVPVRLNQTVIWDRPQVTFSENL 979
Db        796 GSNDENEPETCMVEKMNLTFHVINTGNSMAPNVSVEIMVPNSFS-------PQTDKLFNI 848

Qy        980 ------SSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLK 1033
Db        849 LDVQTTTGECHFENYQ--------------------RVCALEQ--------QKSAMQTLK 880

Qy       1034 GNLSFDWYIKTSHNHLLIVSTAE-----ILFN---------DSVFTLLPGQGAFVRSQTE 1079
Db        881 GIVRF---LSKTDKRLLYCIKADPHCLNFLCNFGKMESGKEASVHIQLEGRPS-ILEMDE 936

Qy       1080 TKVEPFEV-------PNP----------------------------LPLIVGSSVGGLL 1103
Db        937 TSALKFEIRATGFPEPNPRVIELNKDENVAHVLLEGLHHQRPKRYFTIVIISSSLLLGLI 996

Qy       1104 LLALITAALYKLGFFKRQYKDMMSE 1128
Db        997 VLLLISYVMWKAGFFKRQYKSILQE 1021
```

RESULT 14
I58409
integrin alpha-9 chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I58409; A49459
R;Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah
Oncogene 9, 611-619, 1994
A;Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small
A;Reference number: I58409; MUID:94119603; PMID:8290272
A;Accession: I58409
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1035 <RES>
A;Cross-references: UNIPROT:Q13797; GB:D25303; NID:g464180; PIDN:BAA04984.1; PID:g533327
R;Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.
J. Cell Biol. 123, 1289-1297, 1993
A;Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partne
A;Reference number: A49459; MUID:94064789; PMID:8245132
A;Accession: A49459
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 30-1035 <PAL>
A;Cross-references: GB:L24158
C;Superfamily: integrin alpha-4 chain
C;Keywords: glycoprotein; metal binding; transmembrane protein
F;1-27/Domain: signal sequence #status predicted <SIG>

Query Match 10.7%; Score 630; DB 2; Length 1035;
Best Local Similarity 26.5%; Pred. No. 7.5e-35;
Matches 237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;

```
Qy        331 EHEMSQEGFSAAITSNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMT-RVDSDMNDAYLG 389
Db        190 EHGSCQAGIAGFFTEELVVMGAPGSFYWAGTIKVLNLTD-NTYLKLNDEVIMNRRYTYLG 248

Qy        390 YA--AAIILRNRVQSLVLGAPRYQHIGLVAMFR--QNTGMWESNANVKGTQIGAYFGASL 445
Db        249 YAVTAGHFSHPSTIDVVGGAPQDKGIGKVYIFRADRRSGTLIKIFQASGKKMGSYFGSSL 308

Qy        446 CSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPLPRGQRARWQCDAVLYGEQGQPWGRF 505
Db        309 CAVDLNGDGLSDL-LVGAPMFSEIRDEGQVTVY-INRGNGALEE-QLALTGD-GAYNAHF 364

Qy        506 GAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRL 565
Db        365 GESIASLDDLDNDGFPDVAIGAPKEDDFAGAVYIYHGDAG-GIVPQYSMKLSGQKINPVL 423

Qy        566 QYFGQSLSGGQDLTMDGLVDLTVGA--QGHVLLLRSQPVLRVKAIMEFNPREVARNVFEC 623
Db        424 RMFGQSISGGIDMDGNGYPDVTVGAFMSDSVVLLRARPVITVD-VSIFLPGSINITAPQC 482

Qy        624 NDQVVKGKEAGEVRV-CLHVQ-------KSTRDRLREGQIQSVVTYDLALDSGRPHSRAV 675
Db        483 HD--------GQQPVNCLNVTTCFSFHGKHVPEEI---GLNYVLMADVAKKEKGQMPRVY 531

Qy        676 F---NETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNFSL----VGTPLS 728
Db        532 FVLLGETMGQVTEKLQLTYMEETCRHYVAHVKRRVQDVISPIVFEAAYSLSEHVTGEEER 591

Qy        729 AFGNLRPVL-----AEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFM--SLD----CL 777
Db        592 ELPPLTPVLRWKKGQKIAQKNQTV---FERNCRSED-CAADLQLQGKLLLSSMDEKTLYL 647

Qy        778 VVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASS 837
Db        648 ALGAVKNISLNISISNLGDDAYDANVSFNVSRELFF------INMWQKEEMGISCELLES 701

Qy        838 TEVSGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNK 897
Db        702 DFL-------KCSVGFPFMRSKSKYEFSVIFDTSHLS--GEEEVLSFIVTAQSG--NTER 750

Qy        898 TE------FQLELPVKYAVYMVVTS--------HGVSTKYLNFTASENTS---RVMQHQY 940
Db        751 SESLHDNTLVLMVPLMHEVDTSITGIMSPTSFVYGESVDAANFIQLDDLECHFQPINITL 810

Qy        941 QVSNLGQRSLPISLVFL-VPVRLN----------QTVIWDRPQVTFSENLSSTCHTKER 988
Db        811 QVYNTGPSTLPGSSVSISFPNRLSSGGAEMFHVQEMVVGQEKGNCSFQKNPTPCIIPQEQ 870

Qy        989 LPSHSDFLAELRKA---------PVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFD 1039
Db        871 ENIFHTIFAFFTKSGRKVLDCEKPGISCLTAHCN-------FSALAKEESRTI------D 917

Qy       1040 WYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEP--------------- 1084
Db        918 IY--------MLLNT-EILKKDSSSVIQ------FMSRAKVKVDPALRVVEIAHGNPEEV 962

Qy       1085 ---FEVPN---PLPLIVG-----SSVGGLLLLALITAALYKLGFFKRQYKDMM 1126
Db        963 TVVFEALHNLEPRGYVVGWIIAISLLVGILIFLLLAVLLWKMGFFRRRYKEII 1015
```

RESULT 15
T31437
integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C;Species: Lytechinus variegatus (variegated urchin)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31437
R;Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A;Description: Alpha SU2, a sea urchin integrin which binds laminin.
A;Reference number: Z21035
A;Accession: T31437

```
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1041 <HER>
A;Cross-references: UNIPROT:O76378; EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC2
A;Experimental source: developmental stage embryo
C;Function:
A;Description: binds laminin
C;Superfamily: integrin alpha-2b chain

  Query Match           10.5%;  Score 614.5;  DB 2;  Length 1041;
  Best Local Similarity  24.8%;  Pred. No. 8.8e-34;
  Matches  236;  Conservative  155;  Mismatches  314;  Indels  245;  Gaps   41;

Qy        336 QEGFSAAITSNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDS----DMNDAYLG 389
              | |||  | |:    ::   |||   | ::: :  :| :  |: :      :::| |
Db        181 QAGFSGIIFSDNSALVMGAPGSYYLQGQIYVQSLLNRSV-VQATQESNTGTYSFDNSYRG 239

Qy        390 YAAAI--ILRNRVQSLVLGAPRYQHI-GLVAMFRQNTGMWESNANVKGTQIGAYFGASLC 446
              |: |:    : ||  |:| || : : ||||:| ||   :     | |||| |||| |:
Db        240 YSLALGDFNGDGVQDYVVGTPRAESLMGLVAIFDQNLNQFN---QVMGTQIVAYFGYSVT 296

Qy        447 SVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPLPRGQRARWQCDAVLYGEQ-------- 498
              ||: :| : | :|:||| | :      |              ||:  ||    |
Db        297 VVDI-NNDTYDDLLVGAPMYMDDGPAIQ-----------RWEAGAVYVYLQNPDVGPGA 343

Qy        499 ----------GQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGS 546
                       ||   ||| ::  :|| | |   |||||| | |: ||||::|| | :
Db        344 SNRLSLSSTLIGGQIRSRFGLSIASIGDSNQDGFNDVAIGAPYEGDDAGAVYIYHG-SAN 402

Qy        547 GISPSHSQRIAGSKLS-PRLQYFGQSLSGGQDLTMDGLVDLTVGAQ--GHVLLLRSQPVL 603
              |:  : :| :  | |    :  || || ||||:  :   || |||:     :|:|::||:
Db        403 GLKSTPAQVLTPSTLGHSGITTFGFSLQGGQDMDKNKYPDLLVGAESANTAVLIRTRPVV 462

Qy        604 RVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSV----- 658
              : | :   |                    : ::::  | : | :| : :
Db        463 SLDATLNTEP----------------------IGINLENKTYE-LADGTMVTSFIAMT 497

Qy        659 --------------VTYDLALDSG-RPHSRAVFNETKNSTRRQTQVLGL-TQTCETLKLQ 702
                         ::| : :|||   : ||:| :   |   :|: | : || |: |:
Db        498 CFTYTGNYLPDHIDISYTVTVDSGIIANRRAMFVDNDMSEITKTRRLAVSTQFCDPLRAY 557

Qy        703 LPNCIEDPVSPIVLRLNFSLVGTPLSAFGN---------LRPVLAEDAQRLFTALFPFEK 753
              : | ||| ::|| : | : |          |          : |::         |     :
Db        558 VGNSIEDKLTPIKVTLQYDL--------NNDESRLQPHEILPIIDMATMSTQTKQVSIQN 609

Qy        754 NCGNDNICQDDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSY 813
              || | |||  || :| :  :|  :|:|  :|  : |:: | |||:::: :: ::|| | :
Db        610 NCVN-NICIPDLDVTVT-PNLPNIVIGQTQELTLDVSLNNRGEDAFQSSLSVYYPLGLQF 667

Qy        814 RKVSTLQNQRSQRSWRLACESASSTEVSGALKSTSCSINHPIFPEN------SEVTFNIT 867
              ::    :::   : : |   |          |:  :|   :|:  :|       :  || ::
Db        668 VRL----ERKANMDFSVTCSEDSD------LRIITCDTGNPMVGKNILEFGLTLSTFQVS 717

Qy        868 FDVDS-----KASLGNKLLLKANVTSENNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTK 922
              | ||     ||  |      |    | :  |         | :  | |  :  :
Db        718 GDKDSIEFYFKAESENS--EDPNTLENNELNMTVPVTVDCTLKLLSASYPEIVMYSTQED 775

Qy        923 YL-------NFTASENTSRVMQHQYQV-----SNLGQRSLPIS----------LVFLVPV 960
              |:       | : ::    || | |:|     || |: || |           | :|: :
Db        776 YVVPPFPAKNASEADIGMEVM-HLYEVRNTGSSNAGEVSLNIQWPQKNEDGEYLFYLLGI 834

Qy        961 RLNQTVIWDRPQ--------------VTFSENLSSTCHTKERLPSHSDFLAELRKAPVV 1005
              :|          |              | : :     | | |  :: ||:     |:
Db        835 MTEEGVTCQLTQGKANPEGVKLEPSTKAKLSNSTTQVSGRKRREPEVAEALAQTDN--VI 892

Qy       1006 NCSIAVCQRIQCDIPFFGIQEEFNAT------LKGNLSFDWYIKTSHNHLLIVSTAEILF 1059
              |:   |  | | |      :| ||:      : |     |               |  |
Db        893 YCASDSCVLINCTI------DEINASKSKVVRILGRF---W--------------ERTF 928

Qy       1060 NDSVFTLLPGQGAFVRSQTETKVE--PFEVPNP-----------------LP------- 1092
                :|  | |   | : |     |:  |: :| |                 :|
Db        929 QKAVSELTPVVQATIASSASAAVKTIPYNIPLPRDFSDSTKASTLVTTEELVPPVTPIAW 988

Qy       1093 -LIVGSSVGGLLLLALITAALYKLGFFKR-------QYKDMMS--EGGPP 1132
               :|| | :||::|| :|   |:| |||:|       :|  : |  : |||
Db        989 WIIVVSVLGGIILLLIIILGLWKCGFFERKKPGEEKEYAPVASADKDGPP 1038


Search completed: November  9, 2004, 12:19:40
Job time : 33.25 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:46:42 ; Search time 136.75 Seconds
(without alignments)
4783.919 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5852.5	99.6	1152	1	ITAM_HUMAN	P11215 homo sapien
2	4476	76.2	1153	1	ITAM_MOUSE	P05555 mus musculu
3	4387	74.7	1151	2	Q9JI30	Q9ji30 rattus norv
4	3815.5	64.9	920	2	Q28984	Q28984 sus scrofa
5	3483	59.3	1163	1	ITAX_HUMAN	P20702 homo sapien
6	3417	58.2	1162	1	ITAD_HUMAN	Q13349 homo sapien
7	3335.5	56.8	1188	2	Q6KAS4	Q6kas4 mus musculu
8	3335.5	56.8	1188	2	BAD21383	Bad21383 mus muscu
9	3326.5	56.6	1169	1	ITAX_MOUSE	Q9qxh4 mus musculu
10	3251.5	55.3	1161	1	ITAD_RAT	Q9qye7 rattus norv
11	1578	26.9	1165	1	ITAL_BOVIN	P61625 bos taurus
12	1552.5	26.4	1170	1	ITAL_HUMAN	P20701 homo sapien
13	1552	26.4	1166	2	Q6TYB8	Q6tyb8 bos taurus
14	1552	26.4	1166	2	AAQ90015	Aaq90015 bos tauru
15	1533.5	26.1	1163	1	ITAL_MOUSE	P24063 mus musculu
16	1529.5	26.0	1161	2	Q9WTV4	Q9wtv4 mus musculu
17	1519	25.9	1160	2	Q9R200	Q9r200 mus musculu
18	1414	24.1	1196	2	Q98TF1	Q98tf1 cyprinus ca
19	1363.5	23.2	1187	2	Q98TF0	Q98tf0 cyprinus ca
20	1350.5	23.0	1086	2	Q96HB1	Q96hb1 homo sapien
21	1270	21.6	927	2	Q8HZV0	Q8hzv0 bos taurus
22	1163.5	19.8	1167	2	O88340	O88340 rattus norv
23	1157.5	19.7	1167	1	ITAE_MOUSE	Q60677 mus musculu
24	1149	19.6	1179	1	ITAE_HUMAN	P38570 homo sapien
25	1119	19.0	1167	2	O88341	O88341 rattus norv
26	1102.5	18.8	1151	1	ITA1_HUMAN	P56199 homo sapien
27	1093.5	18.6	1189	1	ITAH_HUMAN	Q9ukx5 homo sapien
28	1084	18.5	1170	1	ITA2_BOVIN	P53710 bos taurus
29	1072	18.2	1178	1	ITA2_MOUSE	Q62469 mus musculu
30	1072	18.2	1178	2	Q6P1C7	Q6p1c7 mus musculu
31	1072	18.2	1178	2	AAH65139	Aah65139 mus muscu
32	1069	18.2	1181	1	ITA2_HUMAN	P17301 homo sapien
33	1065.5	18.1	1167	1	ITAG_HUMAN	O75578 homo sapien
34	1060	18.0	1180	1	ITA1_RAT	P18614 rattus norv
35	1052.5	17.9	1188	1	ITAH_MOUSE	P61622 mus musculu
36	1052.5	17.9	1188	2	Q7TQC3	Q7tqc3 mus musculu
37	1051	17.9	1171	2	O42094	O42094 gallus gall
38	1041	17.7	1038	2	Q8BS01	Q8bs01 mus musculu
39	1014	17.3	895	2	Q9WUF8	Q9wuf8 mus sp. itg
40	1013.5	17.3	1160	2	Q8MKF4	Q8mkf4 felis silve
41	865	14.7	348	2	Q8TES5	Q8tes5 homo sapien
42	852	14.5	1332	2	Q9BPQ8	Q9bpq8 halocynthia
43	807	13.7	205	2	Q63001	Q63001 rattus norv
44	784.5	13.4	304	2	Q6PG66	Q6pg66 mus musculu
45	784.5	13.4	304	2	AAH57200	Aah57200 mus muscu

ALIGNMENTS

RESULT 1
ITAM_HUMAN
ID ITAM_HUMAN STANDARD; PRT; 1152 AA.
AC P11215;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha
DE subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)
DE (Neutrophil adherence receptor).
GN Name=ITGAM; Synonyms=CR3A, CD11B;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88315033; PubMed=2457584;
RA Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;
RT "The human leukocyte adhesion glycoprotein Mac-1 (complement receptor
RT type 3, CD11b) alpha subunit. Cloning, primary structure, and relation
RT to the integrins, von Willebrand factor and factor B.";
RL J. Biol. Chem. 263:12403-12411(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88190151; PubMed=2833753;
RA Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;
RT "Molecular cloning of the alpha subunit of human and guinea pig
RT leukocyte adhesion glycoprotein Mo1: chromosomal localization and
RT homology to the alpha subunits of integrins.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88257215; PubMed=2454931;
RA Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;
RT "Amino acid sequence of the alpha subunit of human leukocyte adhesion
RT receptor Mo1 (complement receptor type 3).";
RL J. Cell Biol. 106:2153-2158(1988).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93123748; PubMed=8419480;
RA Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;
RT "Structural analysis of the CD11b gene and phylogenetic analysis of
RT the alpha-integrin gene family demonstrate remarkable conservation of
RT genomic organization and suggest early diversification during
RT evolution.";
RL J. Immunol. 150:480-490(1993).
RN [5]
RP SEQUENCE OF 9-1153 FROM N.A.
RX MEDLINE=89098893; PubMed=2563162;
RA Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L.,
RA Roth G.J.;
RT "cDNA sequence for the alpha M subunit of the human neutrophil
RT adherence receptor indicates homology to integrin alpha subunits.";

```
RL   Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).
RN   [6]
RP   SEQUENCE OF 1-9 FROM N.A.
RX   MEDLINE=92073318; PubMed=1683702;
RA   Shelley C.S., Arnaout M.A.;
RT   "The promoter of the CD11b gene directs myeloid-specific and
RT   developmentally regulated expression.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN   [7]
RP   SEQUENCE OF 1-9 FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=92144986; PubMed=1346576;
RA   Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT   "Characterization of the myeloid-specific CD11b promoter.";
RL   Blood 79:865-870(1992).
RN   [8]
RP   SEQUENCE OF 17-31.
RX   MEDLINE=87076671; PubMed=3539202;
RA   Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
RT   "N-terminal sequence of human leukocyte glycoprotein Mo1: conservation
RT   across species and homology to platelet IIb/IIIa.";
RL   Biochim. Biophys. Acta 874:368-371(1986).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX   MEDLINE=95171458; PubMed=7867070;
RA   Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT   "Crystal structure of the A domain from the alpha subunit of integrin
RT   CR3 (CD11b/CD18).";
RL   Cell 80:631-638(1995).
RN   [10]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX   MEDLINE=96363671; PubMed=8747460;
RA   Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
RT   "Two conformations of the integrin A-domain (I-domain): a pathway for
RT   activation?";
RL   Structure 3:1333-1340(1995).
RN   [11]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX   MEDLINE=98362595; PubMed=9687375;
RA   Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA   Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
RA   Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
RA   Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT   "Cation binding to the integrin CD11b I domain and activation model
RT   assessment.";
RL   Structure 6:923-935(1998).
RN   [12]
RP   3D-STRUCTURE MODELING OF 17-616.
RX   MEDLINE=98226734; PubMed=9560195;
RA   Oxvig C., Springer T.A.;
RT   "Experimental support for a beta-propeller domain in integrin alpha-
RT   subunits and a calcium binding site on its lower surface.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC   -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various
CC       adhesive interactions of monocytes, macrophages and granulocytes
CC       as well as in mediating the uptake of complement-coated particles.
CC       It is identical with CR-3, the receptor for the iC3b fragment of
CC       the third complement component. It probably recognizes the R-G-D
CC       peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for
CC       fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides
CC       of fibrinogen gamma chain.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-M
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC       granulocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J03925; AAA59544.1; -.
DR   EMBL; M18044; AAA59491.1; -.
DR   EMBL; J04145; AAA59903.1; -.
DR   EMBL; S52227; AAB24821.1; -.
DR   EMBL; S52152; AAB24821.1; JOINED.
DR   EMBL; S52153; AAB24821.1; JOINED.
DR   EMBL; S52154; AAB24821.1; JOINED.
DR   EMBL; S52155; AAB24821.1; JOINED.
DR   EMBL; S52157; AAB24821.1; JOINED.
DR   EMBL; S52159; AAB24821.1; JOINED.
DR   EMBL; S52161; AAB24821.1; JOINED.
DR   EMBL; S52164; AAB24821.1; JOINED.
DR   EMBL; S52165; AAB24821.1; JOINED.
DR   EMBL; S52167; AAB24821.1; JOINED.
DR   EMBL; S52169; AAB24821.1; JOINED.
DR   EMBL; S52170; AAB24821.1; JOINED.
DR   EMBL; S52173; AAB24821.1; JOINED.
DR   EMBL; S52174; AAB24821.1; JOINED.
DR   EMBL; S52180; AAB24821.1; JOINED.
DR   EMBL; S52181; AAB24821.1; JOINED.
DR   EMBL; S52184; AAB24821.1; JOINED.
DR   EMBL; S52189; AAB24821.1; JOINED.
DR   EMBL; S52191; AAB24821.1; JOINED.
DR   EMBL; S52192; AAB24821.1; JOINED.
DR   EMBL; S52203; AAB24821.1; JOINED.
DR   EMBL; S52212; AAB24821.1; JOINED.
DR   EMBL; S52213; AAB24821.1; JOINED.
DR   EMBL; S52216; AAB24821.1; JOINED.
DR   EMBL; S52219; AAB24821.1; JOINED.
DR   EMBL; S52220; AAB24821.1; JOINED.
DR   EMBL; S52221; AAB24821.1; JOINED.
DR   EMBL; S52222; AAB24821.1; JOINED.
DR   EMBL; S52226; AAB24821.1; JOINED.
DR   EMBL; M76724; AAA58410.1; -.
DR   EMBL; M84477; AAA51960.1; -.
DR   PIR; A31108; RWHU1B.
DR   PDB; 1A8X; Model; @=17-1152.
DR   PDB; 1BHO; X-ray; 1/2=-.
DR   PDB; 1BHQ; X-ray; 1/2=-.
DR   PDB; 1IDN; X-ray; 1/2=-.
DR   PDB; 1IDO; X-ray; @=140-331.
DR   PDB; 1JLM; X-ray; @=143-334.
DR   PDB; 1M1U; X-ray; A=137-331.
DR   PDB; 1MF7; X-ray; A=144-337.
DR   PDB; 1N9Z; X-ray; A=140-335.
DR   PDB; 1NA5; X-ray; A=144-345.
DR   Genew; HGNC:6149; ITGAM.
DR   MIM; 120980; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0007155; P:cell adhesion; TAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
KW   Glycoprotein; Integrin; Magnesium; Receptor; Repeat; Signal;
KW   Transmembrane.
```

```
FT   SIGNAL        1     16
FT   CHAIN        17   1152       Integrin alpha-M.

  Query Match           99.6%;  Score 5852.5;  DB 1;  Length 1152;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1132;  Conservative    3;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              ||||||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 2
ITAM_MOUSE
ID   ITAM_MOUSE     STANDARD;      PRT;  1153 AA.
AC   P05555; Q8CA73;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha
DE   subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).
GN   Name=Itgam;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=88312584; PubMed=3044779;
RA   Pytela R.;
RT   "Amino acid sequence of the murine Mac-1 alpha chain reveals homology
RT   with the integrin family and an additional domain related to von
RT   Willebrand factor.";
RL   EMBO J. 7:1371-1378(1988).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [3]
RP   SEQUENCE OF 11-45 FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Spleen;
RX   MEDLINE=86287312; PubMed=2942940;
RA   Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E.,
RA   Larson R.S., Roberts T.M., Springer T.A.;
RT   "A partial genomic DNA clone for the alpha subunit of the mouse
```

```
RT   complement receptor type 3 and cellular adhesion molecule Mac-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).
RN   [4]
RP   SEQUENCE OF 17-28.
RX   MEDLINE=85188276; PubMed=3887182;
RA   Springer T.A., Teplow D.B., Dreyer W.J.;
RT   "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
RT   glycoproteins and unexpected relation to leukocyte interferon.";
RL   Nature 314:540-542(1985).
CC   -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various
CC       adhesive interactions of monocytes, macrophages and granulocytes
CC       as well as in mediating the uptake of complement-coated particles.
CC       It is identical with CR-3, the receptor for the iC3b fragment of
CC       the third complement component. It probably recognizes the R-G-D
CC       peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for
CC       fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides
CC       of fibrinogen gamma chain. Alpha-M/beta-2 play a critical role in
CC       mast cell development and in immune complex-mediated
CC       glomerulonephritis. Mice expressing a null mutation of the alpha-M
CC       subunit gene demonstrate increase in neutrophil accumulation, in
CC       response to a impaired degranulation and phagocytosis, events that
CC       apparently accelerate apoptosis in neutrophils. These mice develop
CC       obesity.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta chain. Alpha-M
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P05555-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P05555-2; Sequence=VSP_010473;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC       granulocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X07640; CAA30479.1; -.
DR   EMBL; AK039444; BAC30350.1; -.
DR   EMBL; M14293; AAA39484.1; -.
DR   PIR; S00551; S00551.
DR   HSSP; P11215; 1BHQ.
DR   MGD; MGI:96607; Itgam.
DR   GO; GO:0009897; C:external side of plasma membrane; IDA.
DR   GO; GO:0007155; P:cell adhesion; IMP.
DR   GO; GO:0045123; P:cellular extravasation; IMP.
DR   GO; GO:0030593; P:neutrophil chemotaxis; IMP.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Alternative splicing; Calcium; Cell adhesion;
KW   Direct protein sequencing; Glycoprotein; Integrin; Receptor; Repeat;
KW   Signal; Transmembrane.
```

```
FT   SIGNAL       1    16
FT   CHAIN       17  1153       Integrin alpha-M.
FT   DOMAIN      17  1105       Extracellular (Potential).
FT   TRANSMEM  1106  1129       Potential.
FT   DOMAIN    1130  1153       Cytoplasmic (Potential).
FT   REPEAT      31    84       FG-GAP 1.
FT   REPEAT      85   163       FG-GAP 2.
FT   DOMAIN     164   350       VWFA.
FT   REPEAT     337   400       FG-GAP 3.
FT   REPEAT     401   452       FG-GAP 4.
FT   REPEAT     454   515       FG-GAP 5.
FT   REPEAT     517   575       FG-GAP 6.
FT   REPEAT     580   632       FG-GAP 7.
FT   CA_BIND    465   473       Potential.
FT   CA_BIND    529   537       Potential.
FT   CA_BIND    592   600       Potential.
FT   SITE      1132  1136       GFFKR motif.
FT   DISULFID    66    73       By similarity.
FT   DISULFID   105   123       By similarity.
FT   DISULFID   654   711       By similarity.
FT   DISULFID   770   776       By similarity.
FT   DISULFID   999  1023       By similarity.
FT   DISULFID  1028  1033       By similarity.
FT   CARBOHYD    58    58       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    86    86       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   391   391       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   696   696       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   734   734       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   772   772       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   801   801       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   881   881       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   907   907       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   941   941       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   980   980       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   994   994       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1022  1022       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1045  1045       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1051  1051       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1076  1076       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC   453   569       Missing (in isoform 2).
FT                              /FTId=VSP_010473.
FT   CONFLICT    37    37       N -> S (in Ref. 2).
FT   CONFLICT   683   683       V -> G (in Ref. 2).
SQ   SEQUENCE  1153 AA;  127480 MW;  178DB988AECB0343 CRC64;
```

```
  Query Match           76.2%;  Score 4476;  DB 1;  Length 1153;
  Best Local Similarity  74.3%;  Pred. No. 2.4e-287;
  Matches  845;  Conservative  142;  Mismatches  149;  Indels    2;  Gaps    2;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||: ||||||||:||||:|||| |: ||| |||| | || |:||||||||  | ||
Db         17 FNLDTEHPMTFQENAKGFGQNVVQLGGTSVVVAAPQEAKAVNQTGALYQCDYSTSRCHPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               |||| |||||||||||| :| | |||||||||||| | ||||| |||:|||||| : ||:
Db         77 PLQVPPEAVNMSLGLSLAVSTVPQQLLACGPTVHQNCKENTYVNGLCYLFGSNLLRPPQQ 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              |||||| |||::||| |||||||||    ||::|||||||||||| ||||||||||||:||
Db        137 FPEALRECPQQESDIVFLIDGSGSINNIDFQKMKEFVSTVMEQFKKSKTLFSLMQYSDEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |||||| :|: ||:||| | || || ||| ||:|:|||:||| : |||||:|| |||:||
Db        197 RIHFTFNDFKRNPSPRSHVSPIKQLNGRTKTASGIRKVVRELFHKTNGARENAAKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              |||||||||| |:|||||||| ||||||||||:||   :||:||:|||||  :|||||:
Db        257 TDGEKFGDPLDYKDVIPEADRAGVIRYVIGVGNAFNKPQSRRELDTIASKPAGEHVFQVD 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||| |||||:|||||||||||||:|||||||||||||:|||||||| :|||:||||
Db        317 NFEALNTIQNQLQEKIFAIEGTQTGSTSSFEHEMSQEGFSASITSNGPLLGSVGSFDWAG 376
```

```
Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        377 GAFLYTSKDKVTFINTTRVDSDMNDAYLGYASAVILRNRVQSLVLGAPRYQHIGLVVMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        437 ENFGTWEPHTSIKGSQIGSYFGASLCSVDMDADGNTNLILIGAPHYYEKTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        497 PRG-RARWQCEALLHGDQGHPWGRFGAALTVLGDVNGDKLTDVAIGAPGEQENQGAVYIF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        556 YGASIASLSASHSHRIIGAHFSPGLQYFGQSLSGGKDLTMDGLMDLAVGAQGHLLLLRAQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        616 PVLRLEATMEFSPKKVARSVFACQEQVLKNKDAGEVRVCLRVRKNTKDRLREGDIQSTVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        676 YDLALDPVRSRIRAFFDETKNNTRRRTQVFGLMQKCETLKLILPDCVDDSVSPIILRLNY 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        736 TLVGEPLRSFGNLRPVLAMDAQRFFTAMFPFEKNCGNDSICQDDLSITMSAMGLDTLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRL-ACESASSTE 839
Db        796 GPQDFNMSVTLRNDGEDSYGTQVTVYYPSGLSYRKDSASQNPLTKKPWFVKPAESSSSSE 855

Qy        840 VSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTE 899
Db        856 GHGALKSTTWNINHPIFPANSEVTFNVTFDVDSHASFGNKLLLKAIVASENNMSRTHKTK 915

Qy        900 FQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        916 FQLELPVKYAIYMIVTSDESSIRYLNFTASEMTSKVIQHQYQFNNLGQRSLPVSVVFWIP 975

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        976 VQINNVTVWDHPQVIFSQNLSSACHTEQKSPPHSNFRDQLERTPVLNCSVAVCKRIQCDL 1035

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1036 PSFNTQEIFNVTLKGNLSFDWYIKTSHGHLLLVSSTEILFNDSAFALLPGQESYVRSKTE 1095

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1096 TKVEPYEVHNPVPLIVGSSIGGLVLLALITAGLYKLGFFKRQYKDMMNEAAPQDAPPQ 1153
```

RESULT 3
Q9JI30
ID Q9JI30 PRELIMINARY; PRT; 1151 AA.
AC Q9JI30;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fathallah D.M. Sr., Zerria K. Jr.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AF268593; AAF81280.1; -.
DR HSSP; P11215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 2.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Cell adhesion; Integrin; Transmembrane.
SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.7%; Score 4387; DB 2; Length 1151;
Best Local Similarity 73.1%; Pred. No. 1.9e-281;
Matches 831; Conservative 146; Mismatches 158; Indels 2; Gaps 2;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         17 FNLDTENPMTFQENASGFGQSVIQLGETRVVVAAPQEVKAVNQTGALYQCDYSTNRCDPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         77 PLQVPPEAVNMSLGLSLAATTVPPQLLACGPTVHQNCKENTYVNGLCYLFGSNLLRKPQQ 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        137 FPEALRGCPQQESNIAFLIDGSGSINTIDFQKMKEFVSTVMDQFQKSKTLFSLMQYSDEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db        197 RTHFTFNDFKRNPDPKSHVRPIRQLNGRTKTASGIRKVVRELFQKINGARDNAAKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db        257 TDGEKFGDPLNYEDVIPEAEEAGIIRYVIGVGNAFHKPQSRRELDTIASKPAGDHVFQVD 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        317 NFEALNTIRNQLQEKIFAIEGTQTGSTSSFEHEMSQEGFSAAITSNGPLLGSVGSFDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        377 GAFLYPSKDKASFINTTRIDSDMNDAYLGYASAVISRNRVQSLVLGAPRYQHIGLVVMFK 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        437 QNFGAWEPHTDIKGSQIGSYFGASLCSVDMDADGNTNLILIGAPHHYEKTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        497 PRG-RARWQCEAILHGDQGHPWGRFGAALTVLGDVNGDKLTDVAIGAPGEQENQGAVYIF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        556 HGASVASISTPHSQRIAGARFSPGLQYFGQSLSGGKDLTMDGLMDLAVGAQGRLLLIRAQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        616 PVLRLEATMEFSPKNLARSVFACQEEVVKNKDAGEVRVCLQVRKNTKDRLREGDIQSTVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        676 YDLALDPGRSVVRAFFDETKNGILRRIRVFGLTQKCETLKLILPDCVDNSVSPIILRLNY 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        736 TLVGEPLRSSRDLRPVLAMEAQRIFTAMFPFEKNCGNDTICQDDLSITVSSTRVNTLVVG 795
```

```
Qy        781 GPREFNVVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        796 DSRDFDVSVTLRNDGEDSYGTKVTCYYPSGLSYRKVSASQNQFSKKPWRVIAE-PSSSEG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 QGVLKSTIWDINHPIFPANSEVTFNVTFDVDSDASLGNKLLLKVVVTSENNVSRTDKTES 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        915 QLELPVKYAIYMVVTSGESSIKYLNFTASEMTSKVIQHQYQFNNLGQRSFPVSVVFWIPV 974

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        975 QINKVTIWDPPQVTFSQNLSSVCRTEQKSPSHSKFQDELERTPVLNCSVAVCKRIQCDIP 1034

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1035 SFNSKEIFNVTLQGNLLFDWYIETSHDHLLLVSTAEILFNDSAFALLPGQETFVKAQTET 1094

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1095 KVEPYTVHNPVPLIVGSSVGGLVLLALITAGLYKLGFFKRQYKDMMNEAGGQDGPPQ 1151


RESULT 4
Q28984
ID   Q28984      PRELIMINARY;      PRT;   920 AA.
AC   Q28984;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CD11b (Fragment).
GN   Name=CD11b;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lee J.-K., Schook L.B., Rutherford M.S.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR   EMBL; U40072; AAB16869.1; -.
DR   GO; GO:0008305; C:integrin complex; IEA.
DR   GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 4.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Cell adhesion; Integrin; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER     920    920
SQ   SEQUENCE   920 AA;  102440 MW;  E96CC51E350DD5AC CRC64;

  Query Match             64.9%;  Score 3815.5;  DB 2;  Length 920;
  Best Local Similarity   79.4%;  Pred. No. 1e-243;
  Matches  731;  Conservative   80;  Mismatches  109;  Indels    1;  Gaps    1;

Qy        118 PQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYS 177
Db          1 PQKFPEALRGCPQQESDIAFLIDGSGSINRLDFQRMKEFVSTVMGQFQKSKTLFALMQYS 60

Qy        178 EEFRIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKIL 237
Db         61 EDFYTHFTFNDFKRNPSPKLLVRPIRQLLGRTHTATGIRKVVRELFHSKSGARENALKIL 120

Qy        238 IVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVF 297
Db        121 VVITDGEKFGDPLGYEDVIPEADRKGVIRYVIGVGDAFNSWKSREELNTIASKPCGDHVF 180

Qy        298 QVNNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYD 357
Db        181 QVNNFEAVKTIQNQLQEKTFAIEGTQTGSTSSFECEMSQEGFSAAITSNGPLLGAVGSFD 240

Qy        358 WAGGVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVA 417
Db        241 WAGGAFLHMPKDRVIFINTTRVDSDMNDAYLGYAVEVILRNQAQSLVLGAPRYQHTGLVV 300

Qy        418 MFRQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSV 477
Db        301 MFKQNSGAWEKNADIKGSQIGSYFGASLCSVDVNRDGSSDLVLIGAPHYYEQTRGGQVSV 360

Qy        478 CPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAV 537
Db        361 CPLPQG-RAKWQCRVILCGEQGHPWSRFGAALTALGDVNGDKLTDVAIGAPGEQDNRGAV 419

Qy        538 YLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLL 597
Db        420 YLFHGTSELGISPSHSQRIAGSQLSPRLQYFGQSLSGGQDLTMDGLMDLAVGAQGHVFLL 479

Qy        598 RSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQS 657
Db        480 RSQPVLRVEASMVFKPREVARNVYECRQQAAKTQIAGEVQVCLQVRKSTWDRLREGDTQS 539

Qy        658 VVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLR 717
Db        540 IITYDLALDPGRPHPRAVFEETKNNTRRQTQTLGLSRKCEHLALWLPDCVEDSVTPIVLR 599

Qy        718 LNFSLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCL 777
Db        600 LNFSLVGKPASSFGNLRPVLAVDAQRLFTALFPPEKNCGNDSICQDDLSITFSFMSLDTL 659

Qy        778 VVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASS 837
Db        660 VVGGPRDLKVTLTVRNQGEDSYRTQVTFFYPSDLSYRKVSTSQNQRSQRSWRLACESDVS 719

Qy        838 TEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNK 897
Db        720 TEESTALKSTSCSINHPIFPDNSEVTFNVTFDVNPDAFLGYKLLLKANVTSENNMPSSNK 779

Qy        898 TEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFL 957
Db        780 TEFQLELPVKYAVYVVVTSLEVSTKYFNFTASEKTRHVIEHQYQFNNLGQRKLPISVVFW 839

Qy        958 VPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQC 1017
Db        840 VPVRLNRVTVWDQPQVTFSQNLSRSCSTEEIGPRHSDFLEKLQKTPVLNCSIAVCQKIQC 899

Qy       1018 DIPFFGIQEEFNATLKGNLSF 1038
Db        900 DIPSFGIQEELKVTLKGNLSF 920


RESULT 5
ITAX_HUMAN
ID   ITAX_HUMAN     STANDARD;      PRT;  1163 AA.
AC   P20702; Q8IVA6;
DT   01-FEB-1991 (Rel. 17, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE   alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
GN   Name=ITGAX; Synonyms=CD11C;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE=88166645; PubMed=3327687;
RA   Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT   "cDNA cloning and complete primary structure of the alpha subunit of a
RT   leukocyte adhesion glycoprotein, p150,95.";
RL   EMBO J. 6:4023-4028(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90153906; PubMed=2303426;
RA   Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT   "Genomic structure of an integrin alpha subunit, the leukocyte p150,95
RT   molecule.";
RL   J. Biol. Chem. 265:2782-2788(1990).
RN   [3]
RP   ERRATUM.
RA   Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL   J. Biol. Chem. 265:12750-12751(1990).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 20-43.
RX   MEDLINE=87167596; PubMed=3549901;
RA   Miller L.J., Wiebe M., Springer T.A.;
RT   "Purification and alpha subunit N-terminal sequences of human Mac-1
RT   and p150,95 leukocyte adhesion proteins.";
RL   J. Immunol. 138:2381-2383(1987).
CC   -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
CC       recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
CC       interaction during inflammatory responses. It is especially
CC       important in monocyte adhesion and chemotaxis.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC       granulocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M81695; AAA59180.1; -.
DR   EMBL; M29165; -; NOT_ANNOTATED_CDS.
DR   EMBL; M29487; AAA51620.1; ALT_SEQ.
DR   EMBL; M29482; AAA51620.1; JOINED.
DR   EMBL; M29483; AAA51620.1; JOINED.
DR   EMBL; M29484; AAA51620.1; JOINED.
DR   EMBL; M29485; AAA51620.1; JOINED.
DR   EMBL; M29486; AAA51620.1; JOINED.
DR   EMBL; BC038237; AAH38237.1; -.
DR   PIR; A36584; RWHU1C.
DR   PDB; 1N3Y; X-ray; A=141-338.
DR   Genew; HGNC:6152; ITGAX.
DR   MIM; 151510; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0004872; F:receptor activity; TAS.
DR   GO; GO:0007155; P:cell adhesion; TAS.
DR   GO; GO:0009887; P:organogenesis; TAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
KW   Glycoprotein; Integrin; Magnesium; Polymorphism; Receptor; Repeat;
KW   Signal; Transmembrane.
FT   SIGNAL       1     19
FT   CHAIN       20   1163       Integrin alpha-X.
FT   DOMAIN      20   1107       Extracellular (Potential).
FT   TRANSMEM  1108   1128       Potential.
FT   DOMAIN    1129   1163       Cytoplasmic (Potential).
FT   REPEAT      34     87       FG-GAP 1.
FT   REPEAT      88      ?       FG-GAP 2.
FT   DOMAIN     165    351       VWFA.
FT   REPEAT       ?    401       FG-GAP 3.
FT   REPEAT     402    453       FG-GAP 4.
FT   REPEAT     455    517       FG-GAP 5.
FT   REPEAT     518    576       FG-GAP 6.
FT   REPEAT     581    633       FG-GAP 7.
FT   CA_BIND    466    474       Potential.
FT   CA_BIND    530    538       Potential.
FT   CA_BIND    593    601       Potential.
FT   SITE      1131   1135       GFFKR motif.
FT   DISULFID    69     76       By similarity.
FT   DISULFID   108    126       By similarity.
FT   DISULFID   655    712       By similarity.
FT   DISULFID   771    777       By similarity.
FT   DISULFID   848    863       By similarity.
FT   DISULFID   998   1022       By similarity.
FT   DISULFID  1027   1032       By similarity.
FT   CARBOHYD    61     61       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   392    392       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   697    697       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   735    735       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   899    899       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   939    939       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1050   1050       N-linked (GlcNAc. . .) (Potential).
FT   VARIANT     48     48       W -> R (in dbSNP:11574633).
FT                               /FTId=VAR_018672.
FT   CONFLICT   209    209       T -> S (in Ref. 4).
FT   CONFLICT   251    251       T -> A (in Ref. 4).
FT   CONFLICT   469    469       T -> S (in Ref. 4).
FT   CONFLICT   490    490       G -> A (in Ref. 2).
FT   CONFLICT   547    547       E -> K (in Ref. 4).
FT   CONFLICT   756    756       D -> L (in Ref. 1).
FT   CONFLICT   819    819       I -> V (in Ref. 4).
FT   CONFLICT  1161   1163       SEK -> TPHYPQDNV (in Ref. 4).
FT   STRAND     150    157
```

```
FT   TURN        160    161
FT   HELIX       164    178
FT   TURN        179    180
FT   TURN        183    185
FT   STRAND      186    193
FT   STRAND      197    201
FT   HELIX       203    208
FT   HELIX       212    216
FT   TURN        217    218
FT   STRAND      226    226
FT   HELIX       228    236
FT   TURN        237    240
FT   HELIX       242    244
FT   TURN        245    245
FT   TURN        248    249
FT   STRAND      251    258
FT   STRAND      263    263
FT   HELIX       269    278
FT   TURN        279    280
FT   STRAND      282    288
FT   HELIX       290    293
FT   TURN        296    297
FT   HELIX       298    304
FT   HELIX       310    312
FT   STRAND      313    316
FT   HELIX       319    325
FT   HELIX       326    334
FT   TURN        335    335
SQ   SEQUENCE   1163 AA;  127886 MW;  83658A13B5C5DE8F CRC64;

  Query Match             59.3%;  Score 3483;  DB 1;  Length 1163;
  Best Local Similarity   61.4%;  Pred. No. 1.6e-221;
  Matches  693;  Conservative  138;  Mismatches  292;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||| |: :: ||| |||| | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||:|||| |||||||||| | | |: ||||| |    |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || ||||||||| :| | || |: | :: | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              : ||||:||:   || ||:  : || | |:||| :: |:  | : : |||::| ||||||
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              |||:| || | |:|||| ||  |:||| |||| ||::  | :||| ||||| ::|:|:|
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:||||||||:| |||||| ||:|||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      ||||:: : || |:||||:  : |   ||||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:||||||||||||::|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  |||||||||||:|| |||||||:|||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :||| :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:|| |||| |||  ||||||| |:||||:|  |:|||   |  |:||
Db        737 TLVGKPLLAFRNLRPMLAADAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |  ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::|| |||| || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||||| ||:||   ||||||: | |  | |  |:|||:||||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | ||| :|   :|:  :| | | ::::  | ||||| ::| ||::|||| | | :|:|:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :     : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143


RESULT 6
ITAD_HUMAN
ID   ITAD_HUMAN     STANDARD;      PRT;  1162 AA.
AC   Q13349; Q15575; Q15576;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN   Name=ITGAD;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=96111956; PubMed=8777714;
RA   Van der Vieren M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA   Staunton D.E., Gallatin W.M.;
RT   "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
RT   3.";
RL   Immunity 3:683-690(1995).
RN   [2]
RP   SEQUENCE OF 1-235 FROM N.A.
RX   MEDLINE=20187620; PubMed=10722744;
RA   Noti J.D., Johnson A.K., Dillon J.D.;
RT   "Structural and functional characterization of the leukocyte integrin
RT   gene CD11d. Essential role of Sp1 and Sp3.";
RL   J. Biol. Chem. 275:8959-8969(2000).
RN   [3]
RP   SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX   MEDLINE=96257236; PubMed=8666289;
RA   Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
RT   "Cloning and chromosomal localization of a novel gene-encoding a human
RT   beta 2-integrin alpha subunit.";
```

```
RL   Gene 171:291-294(1996).
RN   [4]
RP   INTERACTION WITH VCAM1.
RX   MEDLINE=99059842; PubMed=9841932;
RA   Grayson M.H., Van der Vieren M., Sterbinsky S.A., Michael Gallatin W.,
RA   Hoffman P.A., Staunton D.E., Bochner B.S.;
RT   "alphadbeta2 integrin is expressed on human eosinophils and functions
RT   as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-
RT   1).";
RL   J. Exp. Med. 188:2187-2191(1998).
RN   [5]
RP   INTERACTION WITH VCAM1.
RX   MEDLINE=99370002; PubMed=10438935;
RA   Van der Vieren M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
RA   Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RT   "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
RT   binding interface between I domain and VCAM-1.";
RL   J. Immunol. 163:1984-1990(1999).
CC   -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC       VCAM1. May play a role in the atherosclerotic process such as
CC       clearing lipoproteins from plaques and in phagocytosis of blood-
CC       borne pathogens, particulate matter, and senescent erythrocytes
CC       from the blood.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell
CC       lines and subsets of peripheral blood leukocytes and strongly on
CC       tissue-specialized cells, including macrophages foam cells within
CC       atherosclerotic plaques, and on splenic red pulp macrophages.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U37028; AAB38547.1; -.
DR   EMBL; U40274; AAB60634.1; -.
DR   EMBL; U40275; AAB60635.1; -.
DR   EMBL; U40276; AAB60636.1; -.
DR   EMBL; U40277; AAB60637.1; -.
DR   EMBL; U40279; AAB60638.1; -.
DR   EMBL; U40278; AAB60638.1; JOINED.
DR   EMBL; AF187881; AAF62875.1; -.
DR   HSSP; P11215; 1BHQ.
DR   Genew; HGNC:6146; ITGAD.
DR   MIM; 602453; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0016337; P:cell-cell adhesion; NAS.
DR   GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     17       Potential.
FT   CHAIN        18   1162       Integrin alpha-D.
FT   DOMAIN       18   1100       Extracellular (Potential).
FT   TRANSMEM   1101   1121       Potential.
FT   DOMAIN     1122   1162       Cytoplasmic (Potential).
FT   REPEAT       32     85       FG-GAP 1.
FT   REPEAT       86      ?       FG-GAP 2.
FT   DOMAIN      150    332       VWFA.
FT   REPEAT      350    400       FG-GAP 3.
FT   REPEAT      401    452       FG-GAP 4.
FT   REPEAT      454    516       FG-GAP 5.
FT   REPEAT      518    576       FG-GAP 6.
FT   REPEAT      581    633       FG-GAP 7.
FT   CA_BIND     465    473       Potential.
FT   CA_BIND     530    538       Potential.
FT   CA_BIND     593    601       Potential.
FT   SITE       1127   1131       GFFKR motif.
FT   DISULFID     67     74       By similarity.
FT   DISULFID    106    124       By similarity.
FT   DISULFID    655    710       By similarity.
FT   DISULFID    769    775       By similarity.
FT   DISULFID    846    861       By similarity.
FT   DISULFID    994   1018       By similarity.
FT   DISULFID   1023   1028       By similarity.
FT   CARBOHYD      59     59       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD      87     87       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD      99     99       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     391    391       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     691    691       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     733    733       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     873    873       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     957    957       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    1046   1046       N-linked (GlcNAc. . .) (Potential).
FT   CONFLICT     500    500       Missing (in Ref. 2).
FT   CONFLICT     515    518       GHPW -> ATP (in Ref. 2).
FT   CONFLICT     825    825       L -> V (in Ref. 2).
FT   CONFLICT     984    984       V -> A (in Ref. 2).
SQ   SEQUENCE   1162 AA;  126885 MW;  F296A1A35455D77D CRC64;

  Query Match           58.2%;  Score 3417;  DB 1;  Length 1162;
  Best Local Similarity 59.8%;  Pred. No. 3.7e-217;
  Matches  675;  Conservative  162;  Mismatches  284;  Indels    8;  Gaps    6;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    |||:| |||||||| |||:||||| |:||||| | || |  :|| |:||
Db         18 FNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQTGRLYDCAAATGMCQPI 77

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               | :  |||||||||:|||:|:  :|||||||:|: | ||:| || | | ||   :  |
Db         78 PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-EIIQT 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              |:|   || :: || |||||||||  :|| :|| ||  || | : : |||:|||||
Db        137 VPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFEGTDTLFALMQYSNLL 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              :||||| :|: :|: :||| || || | | |||:  |: :| :  |||||:| ||||||
Db        197 KIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHHKNGARKSAKKILIVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              |||:|: ||| | ||||:|:: |:||| |||| ||:   :|||||||:| ||:||||:|:
Db        257 TDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELNTISSAPPQDHVFKVD 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              || || :|| ||:|||:|:||||: :||||:|||||||||| |:| :|  |  |||: |:|
Db        317 NFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGSFSWSG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |   ||:|||||||||| |   :|
Db        377 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVLGAPRYQHTGKAVIFT 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:||||||||||||:|||||:|||||||||||||||||||
Db        437 QVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPHYYEQTRGGQVSVCPL 496
```

```
Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||| :||||||| |||| |||||||||||||||| ||| |||||||||::||||||||
Db        497 PRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAIGAPGEQENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || | |||||||||||| |:||||||||||:|||||||| |||:|| |||:| ||||||
Db        557 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLLRSL 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |||:|   | |:| |||: |: | ::     |||:  ||| :|||: |:|  | |||  |
Db        617 PVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 674

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              :||||| ||  |||:|||||| |  : : |||   |||||| ||:|:|| ||||:| |||
Db        675 FDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCVEDVVSPIILHLNF 734

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||  |: :  |||||||  :| ||||  ||||||| | :|: || :| ||  |  | ||
Db        735 SLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDLGVTLSFSGLQTLTVG 794

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | || ||| | ||||| | |: ::|  ||:|:||  | |  | : |||||:   ||
Db        795 SSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 853

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              | |:|: ||:|||| | |  || :|||  ||:||:::|::|: :|||   ::|  |
Db        854 EG-LRSSRCSVNHPIFHEGSNGTFIVTFDVSYKATLGDRMLMRASASSENNKASSSKATF 912

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNF-TASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||||| :::    |||| || |: |   :  :|:|:|:|:|| || | ||: | ||
Db        913 QLELPVKYAVYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQRDLAISINFWVP 972

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | ||   :||       |::|   | :: : | ||||| :: ::|:::|||| | : :||:
Db        973 VLLNGVAVWDVVMEAPSQSL--PCVSERKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1030

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :|    :|:|| ||| |: ||:: |||| ||:|:| |
Db       1031 PSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQLPGQEAFMRAQME 1090

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
                :|  || | :|:|:||||| ||||||||| ||||||||| ||:|: :
Db       1091 MVLEEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRHYKEMLED 1139
```

RESULT 7
Q6KAS4
ID Q6KAS4 PRELIMINARY; PRT; 1188 AA.
AC Q6KAS4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MFLJ00114 protein (Fragment).
GN Name=mFLJ00114;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Genes:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-Homologous cDNAs
RT Identified by Screening of Terminal Sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:167-180(2004).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AK131133; BAD21383.1; -.
DR GO; GO:0009897; C:external side of plasma membrane; IDA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; UNKNOWN_1.
DR PROSITE; PS50234; VWFA; 1.
KW Cell adhesion; Integrin; Transmembrane.
FT NON_TER 1 1
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDBB4178 CRC64;

Query Match 56.8%; Score 3335.5; DB 2; Length 1188;
Best Local Similarity 57.4%; Pred. No. 9.8e-212;
Matches 654; Conservative 165; Mismatches 302; Indels 19; Gaps 7;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    |  :   || ||:|   | ||||||:|| | || | ||:| | ||:||||
Db         39 FNLDAEKPTHFHMDGAEFGHSVLQYDSSWVVVGAPKEIKATNQIGGLYKCGYHTGNCEPI 98

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               |||| ||||||||||||| |:|  |||||||| || || |: ||||  |: :|  |
Db         99 SLQVPPEAVNMSLGLSLAAATNPSWLLACGPTVHHTCRENIYLTGLCFLLSSSFKQS-QN 157

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              || | : ||:|| || |||||||||   || :| :||  || ||::  | |||||:|: |
Db        158 FPTAQQECPKQDQDIVFLIDGSGSISSTDFEKMLDFVKAVMSQLQRPSTRFSLMQFSDYF 217

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |:|||  | :  :| ||:  : || | |:||: :: || ||    :|||::| |:||||
Db        218 RVHFTFNNFISTSSPLSLLDSVRQLRGYTYTASAIKHVITELFTTQSGARQDATKVLIVI 277

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||| | || | |: ||| |:   :||| |||| || :| |:|||  ||| |  ::|| |
Db        278 TDGRKQGDNLSYDSVIPMAEAASIIRYAIGVGKAFYNEHSKQELKAIASMPSHEYVFSVE 337

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||:||| |:|||:|||||||||:| |||:|| |||||||||  | :||:|  |||: |:|
Db        338 NFDALKDIENQLKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTPDGPVLGAVGSFSWSG 397

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | ||| |  : ||||:: : || |||||: |:     | ||:|||||:|| | | :|
Db        398 GAFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVVIFT 457

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   : |:|||||:|||||||||:| :|||||||| ||||| ||||||||||:
Db        458 QESRHWRPKSEVRGTQIGSYFGASLCSVDMDRDGSTDLVLIGVPHYYEHTRGGQVSVCPM 517

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              | |   || |   |:|||| |||||||||||||||| | ||||||||||:||||||:|
Db        518 P-GVGRRWHCGTTLHGEQGHPWGRFGAALTVLGDVNGDSLADVAIGAPGEEENRGAVYIF 576

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || |   |:|| |||| |::  |:|||||||||||||| |||||| ||::| |||||::
Db        577 HGASRQDIAPSPSQRISASQIPSRIQYFGQSLSGGQDLTRDGLVDLAVGSKGRVLLLRTR 636

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |:|||   : | | |::|:|||| :||   :   ||||: :| : :|  | ::| ||
Db        637 PILRVSPTVHFTPAEISRSVFECQEQVAPEQTLSDATVCLHIHESPKTQL--GDLRSTVT 694

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              :||||| ||  :||:| |||     : : ||| : ||::|| || |:|| |:|| |||||
Db        695 FDLALDHGRLSTRAIFKETKTRALTRVKTLGLNKHCESVKLLLPACVEDSVTPITLRLNF 754

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||| |:|:  ||:|:|| | |  |||  |||||| |:||||||: | |  |  ||||
Db        755 SLVGVPISSLQNLQPMLAVDDQTYFTASLPFEKNCGADHICQDDLSVVFGFPDLKTLVVG 814

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQ----------NQRSQRSWR 829
              | || ||| ||||||| | || |:|: ||:|:|:  |            :| | |
Db        815 SDLELNVDVTVSNDGEDSYGTTVTLFYPVGLSFRRVAEGQVFLRKKEDQQWQRRGQHSLH 874
```

```
Qy        830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSE 889
              | |:  |: : |  | |||||  | ||   |::|| :|||| || ||::|||:| | ||
Db        875 LMCD--STPDRSQGLWSTSCSSRHVIFRGGSQMTFLVTFDVSPKAELGDRLLLRARVGSE 932

Qy        890 NNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASE-NTSRVMQHQYQVSNLGQR 948
              ||:| | || ||||||||||||| :::||   ||||||: ||   : |::|::||:|||||
Db        933 NNVPGTPKTTFQLELPVKYAVYTMISSHDQFTKYLNFSTSEKEKTSVVEHRFQVNNLGQR 992

Qy        949 SLPISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCS 1008
              :|:|: | ||: |   :|   |:  :|  : |:     |:  |||  ::|:||::||
Db        993 DVPVSINFWVPIELKGEAVW-TVMVSHPQNPLTQCYRNRLKPTQFDFLTHMQKSPVLDCS 1051

Qy       1009 IAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLP 1068
              || |  ::|||| || :|   |||||| |  :|   :|::| ||| || ||:: ||
Db       1052 IADCLHLRCDIPSLGILDELYFILKGNLSFGWISQTLQKKVLLLSEAEITFNTSVYSQLP 1111

Qy       1069 GQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              || ||:|:||:| :| ::| ||:|||||||||||||||:||| ||| ||||||||:|: |
Db       1112 GQEAFLRAQTKTVLEMYKVHNPVPLIVGSSVGGLLLLAIITAILYKAGFFKRQYKEMLEE 1171


RESULT 8
BAD21383
ID   BAD21383    PRELIMINARY;      PRT;  1188 AA.
AC   BAD21383;
DT   01-JUN-2004 (TrEMBLrel. 27, Created)
DT   01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT   01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE   MFLJ00114 protein (Fragment).
GN   MFLJ00114.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA   Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT   "Prediction of the Coding Sequences of Mouse Homologues of FLJ Gene:
RT   The Complete Nucleotide Sequences of 110 Mouse FLJ-homologous cDNAs
RT   Identified by Screening of Terminal sequences of cDNA Clones Randomly
RT   Sampled from Size-Fractionated Libraries.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK131133; BAD21383.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   1188 AA;  131248 MW;  B8D93107BDBB4178 CRC64;

  Query Match             56.8%;  Score 3335.5;  DB 2;  Length 1188;
  Best Local Similarity   57.4%;  Pred. No. 9.8e-212;
  Matches  654;  Conservative  165;  Mismatches  302;  Indels   19;  Gaps    7;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    |  :   || ||:|   | |||||||:|| | || | ||:| | ||:||||
Db         39 FNLDAEKPTHFHMDGAEFGHSVLQYDSSWVVVGAPKEIKATNQIGGLYKCGYHTGNCEPI 98

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||||||||||| |:|  |||||||| || || |: |||||  |: :|  |
Db         99 SLQVPPEAVNMSLGLSLAAATNPSWLLACGPTVHHTCRENIYLTGLCFLLSSSFKQS-QN 157

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              || | : ||::| || |||||||||   || :| :||  || ||::  | |||||:|: |
Db        158 FPTAQQECPKQDQDIVFLIDGSGSISSTDFEKMLDFVKAVMSQLQRPSTRFSLMQFSDYF 217

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |:|||  | :  :| ||:  : || | |:||: :: || ||    :|||::| |:||||
Db        218 RVHFTFNNFISTSSPLSLLDSVRQLRGYTYTASAIKHVITELFTTQSGARQDATKVLIVI 277

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||| | || | |: ||| |:   :||| |||| || :| |:|||  ||| |  ::|| |
Db        278 TDGRKQGDNLSYDSVIPMAEAASIIRYAIGVGKAFYNEHSKQELKAIASMPSHEYVFSVE 337

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||:||| |:|||:|||||||||:| |||:|| |||||||||  | :||:|  |||: |:|
Db        338 NFDALKDIENQLKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTPDGPVLGAVGSFSWSG 397

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | ||| |  : |||||:: : || |||||: |:      | ||:|||||:|| | | :|
Db        398 GAFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVVIFT 457

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   : |:|||||:|||||||||:| :|||||||| ||||| ||||||||||:
Db        458 QESRHWRPKSEVRGTQIGSYFGASLCSVDMDRDGSTDLVLIGVPHYYEHTRGGQVSVCPM 517

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              | |   || |   |:|||| |||||||||||||||||| | ||||||||||:||||||:|
Db        518 P-GVGRRWHCGTTLHGEQGHPWGRFGAALTVLGDVNGDSLADVAIGAPGEEENRGAVYIF 576

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || |   |:|| |||||: |::  |:|||||||||||||| ||||| ||::| |||||::
Db        577 HGASRQDIAPSPSQRISASQIPSRIQYFGQSLSGGQDLTRDGLVDLAVGSKGRVLLLRTR 636

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |:|||   : | | |::|:|||| :||   :   :  ||||: :| : :|  | ::| ||
Db        637 PILRVSPTVHFTPAEISRSVFECQEQVAPEQTLSDATVCLHIHESPKTQL--GDLRSTVT 694

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              :|||||  ||  :||:| |||      : : ||| : ||::|| || |:|| |:|| |||||
Db        695 FDLALDHGRLSTRAIFKETKTRALTRVKTLGLNKHCESVKLLLPACVEDSVTPITLRLNF 754

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||| |:|:  ||:|:|| | |  |||  |||||||  |:|||||: | |  |  ||||
Db        755 SLVGVPISSLQNLQPMLAVDDQTYFTASLPPEKNCGADHICQDDLSVVFGFPDLKTLVVG 814

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQ----------NQRSQRSWR 829
              | || ||| ||||||| | || |:|: ||:|:|:  |           :| | |
Db        815 SDLELNVDVTVSNDGEDSYGTTVTLFYPVGLSFRRVAEGQVFLRKKEDQQWQRRGQHSLH 874

Qy        830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSE 889
              | |:  |: : |  | |||||  | ||   |::|| :||||  || ||::|||:| | ||
Db        875 LMCD--STPDRSQGLWSTSCSSRHVIFRGGSQMTFLVTFDVSPKAELGDRLLLRARVGSE 932

Qy        890 NNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASE-NTSRVMQHQYQVSNLGQR 948
              ||:| | || ||||||||||||| :::||   ||||||: ||   : |::|::||:|||||
Db        933 NNVPGTPKTTFQLELPVKYAVYTMISSHDQFTKYLNFSTSEKEKTSVVEHRFQVNNLGQR 992

Qy        949 SLPISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCS 1008
              :|:|: | ||: |   :|   |:  :|  : |:     |:  |||  ::|:||::||
Db        993 DVPVSINFWVPIELKGEAVW-TVMVSHPQNPLTQCYRNRLKPTQFDFLTHMQKSPVLDCS 1051

Qy       1009 IAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLP 1068
              || |  ::|||| || :|   |||||| |  :|   :|::| ||| || ||:: ||
Db       1052 IADCLHLRCDIPSLGILDELYFILKGNLSFGWISQTLQKKVLLLSEAEITFNTSVYSQLP 1111

Qy       1069 GQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              || ||:|:||:| :| ::| ||:|||||||||||||||:||| ||| ||||||||:|: |
Db       1112 GQEAFLRAQTKTVLEMYKVHNPVPLIVGSSVGGLLLLAIITAILYKAGFFKRQYKEMLEE 1171


RESULT 9
ITAX_MOUSE
ID   ITAX_MOUSE     STANDARD;      PRT;  1169 AA.
AC   Q9QXH4;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE   alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c).
GN   Name=Itgax;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Dendritic cell;
```

```
RA   Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RA   Tsuchiya H.;
RT   "Isolation of genes selectively expressed by dendritic cells.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
CC       recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
CC       interaction during inflammatory responses. It is especially
CC       important in monocyte adhesion and chemotaxis (By similarity).
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
CC       associates with beta-2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF211864; AAF23492.1; -.
DR   HSSP; P20702; 1N3Y.
DR   MGD; MGI:96609; Itgax.
DR   GO; GO:0009897; C:external side of plasma membrane; IDA.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     19       Potential.
FT   CHAIN        20   1169       Integrin alpha-X.
FT   DOMAIN       20   1116       Extracellular (Potential).
FT   TRANSMEM   1117   1137       Potential.
FT   DOMAIN     1138   1169       Cytoplasmic (Potential).
FT   REPEAT       34     87       FG-GAP 1.
FT   REPEAT       88      ?       FG-GAP 2.
FT   DOMAIN      152    330       VWFA.
FT   REPEAT        ?    402       FG-GAP 3.
FT   REPEAT      403    454       FG-GAP 4.
FT   REPEAT      456    518       FG-GAP 5.
FT   REPEAT      519    577       FG-GAP 6.
FT   REPEAT      582    634       FG-GAP 7.
FT   CA_BIND     467    475       Potential.
FT   CA_BIND     531    539       Potential.
FT   CA_BIND     594    602       Potential.
FT   SITE       1140   1144       GFFKR motif.
FT   DISULFID     69     76       By similarity.
FT   DISULFID    108    126       By similarity.
FT   DISULFID    656    711       By similarity.
FT   DISULFID    770    776       By similarity.
FT   DISULFID    858    873       By similarity.
FT   DISULFID   1007   1031       By similarity.
FT   DISULFID   1036   1041       By similarity.
FT   CARBOHYD     89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    267    267       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    393    393       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    734    734       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    949    949       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1059   1059       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1084   1084       N-linked (GlcNAc. . .) (Potential).
```

```
SQ   SEQUENCE   1169 AA;  129150 MW;  C616412033C219A6 CRC64;

  Query Match             56.6%;  Score 3326.5;  DB 1;  Length 1169;
  Best Local Similarity   57.2%;  Pred. No. 3.8e-211;
  Matches  652;  Conservative  167;  Mismatches  302;  Indels   19;  Gaps    7;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    |  :   || ||:|   | ||||||:|| | || | ||:| | ||:||||
Db         20 FNLDAEKLTHFHMDGAEFGHSVLQYDSSWVVVGAPKEIKATNQIGGLYKCGYHTGNCEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||:|||||||| |:|  |||||||| || || |: |||||  |: :|  |
Db         80 SLQVPPEAVNISLGLSLAAATNPSWLLACGPTVHHTCRENIYLTGLCFLLSSSFKQS-QN 138

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              || | : ||::| || |||||||||   || :| :||  || ||::  | |||||:|: |
Db        139 FPTAQQECPKQDQDIVFLIDGSGSISSTDFEKMLDFVKAVMSQLQRPSTRFSLMQFSDYF 198

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |:|||  | :  :| ||:  : || | |:||: :: || ||    :|||::| |:|||
Db        199 RVHFTFNNFISTSSPLSLLGSVRQLRGYTYTASAIKHVITELFTTQSGARQDATKVLIVI 258

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||| | || | |: ||| |:   :||| |||| || :| |:|||  ||| |  ::|| |
Db        259 TDGRKQGDNLSYDSVIPMAEAASIIRYAIGVGKAFYNEHSKQELKAIASMPSHEYVFSVE 318

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||:||| |:|||:|||||||||:| |||:|| |||||||||  | :||:|  |||: |:|
Db        319 NFDALKDIENQLKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTPDGPVLGAVGSFSWSG 378

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | ||| |  : |||||:: : || ||||||: |:      | ||:||||:|| | | :|
Db        379 GAFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVVIFT 438

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   : |:|||||:||||||||||:| :||||||||| ||||| ||||||||||:
Db        439 QESRHWRPKSEVRGTQIGSYFGASLCSVDMDRDGSTDLVLIGVPHYYEHTRGGQVSVCPM 498

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              | |  :|| |   |:|||| |||||||||||||||||| | ||||||||||:|||||:|
Db        499 P-GVGSRWHCGTTLHGEQGHPWGRFGAALTVLGDVNGDSLADVAIGAPGEEENRGAVYIF 557

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || |   |:|| |||| |::  |:|||||||||||| ||||| ||::| |||| ::
Db        558 HGASRQDIAPSPSQRISASQIPSRIQYFGQSLSGGQDLTRDGLVDLAVGSKGRVLLLRTR 617

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |:|||   : | | |::|:|||| :||   :   :  ||||: :| : :|  | ::| ||
Db        618 PILRVSPTVHFTPAEISRSVFECQEQVAPEQTLSDATVCLHIHESPKTQL--GDLRSTVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              :|||| ||  :||:| |||     : : ||| : ||::|| || |:|| |:|| |||||
Db        676 FDLALDHGRLSTRAIFKETKTRALTRVKTLGLNKHCESVKLLLPACVEDSVTPITLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||| |:|:  ||:|:|| | |  ||| |||||| |:|||||: | |  |  ||||
Db        736 SLVGVPISSLQNLQPMLAVDDQTYFTASLPFEKNCGADHICQDDLSVVFGFPDLKTLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQ-----------NQRSQRSWR 829
                | || ||| ||||||| | || |:|: ||:|:|:  |           :| | |
Db        796 SDLELNVDVTVSNDGEDSYGTTVTLFYPVGLSFRRVAEGQVFLRKKEDQQWQRRGQHSLH 855

Qy        830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSE 889
              | |:  |: : |  | |||||  | ||   |::|| :|||| || ||::|||:| | ||
Db        856 LMCD--STPDRSQGLWSTSCSSRHVIFRGGSQMTFLVTFDVSPKAELGDRLLLRARVGSE 913

Qy        890 NNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASE-NTSRVMQHQYQVSNLGQR 948
              ||:| | || |||||||||||| :::||   ||||||: ||   :  |::|::||:|||||
Db        914 NNVPGTPKTTFQLELPVKYAVYTMISSHDQFTKYLNFSTSEKEKTSVVEHRFQVNNLGQR 973

Qy        949 SLPISLVFLVPVRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCS 1008
              :|:|: | ||: |   :|   |:  :|  : |:      |:  | |  ::|:||::||
Db        974 DVPVSINFWVPIELKGEAVW-TVMVSHPQNPLTQCYRNRLKPTQFDLLTHMQKSPVLDCS 1032
```

```
Qy       1009 IAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLP 1068
              || |  ::|||| || :|   |||||| |  :|   :|::| ||| || ||:: ||
Db       1033 IADCLHLRCDIPSLGILDELYFILKGNLSFGWISQTLQKKVLLLSEAEITFNTSVYSQLP 1092

Qy       1069 GQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              || ||:|:||:| :| ::| ||:|||||||||||||:||| ||| ||||||||:|: |
Db       1093 GQEAFLRAQTKTVLEMYKVHNPVPLIVGSSVGGLLLLAIITAILYKAGFFKRQYKEMLEE 1152


RESULT 10
ITAD_RAT
ID   ITAD_RAT       STANDARD;      PRT;  1161 AA.
AC   Q9QYE7;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-D precursor.
GN   Name=Itgad;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RA   O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA   Gallatin W.M.;
RT   "Cloning of rat alpha D, a novel beta 2 integrin.";
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC       VCAM1. May play a role in the atherosclerotic process such as
CC       clearing lipoproteins from plaques and in phagocytosis of blood-
CC       borne pathogens, particulate matter, and senescent erythrocytes
CC       from the blood (By similarity).
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC       associates with beta-2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF021334; AAF21241.1; -.
DR   HSSP; P11215; 1BHQ.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     19       Potential.
FT   CHAIN        20   1161       Integrin alpha-D.
FT   DOMAIN       20   1100       Extracellular (Potential).
FT   TRANSMEM   1101   1121       Potential.
FT   DOMAIN     1122   1161       Cytoplasmic (Potential).
FT   REPEAT       34     87       FG-GAP 1.
FT   REPEAT       88      ?       FG-GAP 2.
FT   DOMAIN      152    334       VWFA.
FT   REPEAT      352    402       FG-GAP 3.
FT   REPEAT      403    454       FG-GAP 4.
FT   REPEAT      456    517       FG-GAP 5.
FT   REPEAT      519    577       FG-GAP 6.
FT   REPEAT      582    634       FG-GAP 7.
FT   CA_BIND     467    475       Potential.
FT   CA_BIND     531    539       Potential.
FT   CA_BIND     594    602       Potential.
FT   SITE       1126   1130       GFFKR motif.
FT   DISULFID     69     76       By similarity.
FT   DISULFID    108    126       By similarity.
FT   DISULFID    656    711       By similarity.
FT   DISULFID    769    775       By similarity.
FT   DISULFID    845    860       By similarity.
FT   DISULFID    993   1017       By similarity.
FT   DISULFID   1022   1027       By similarity.
FT   CARBOHYD     61     61       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    245    245       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    393    393       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    696    696       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    734    734       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    784    784       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    907    907       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    936    936       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1045   1045       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   1161 AA;  126600 MW;  2258491A9B4A705E CRC64;

  Query Match             55.3%;  Score 3251.5;  DB 1;  Length 1161;
  Best Local Similarity   58.0%;  Pred. No. 3.5e-206;
  Matches  654;  Conservative  158;  Mismatches  302;  Indels   13;  Gaps    9;

Qy          2 NLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPIR 61
              ||| |  : |:|:|  |||:||| |||:||||| | || || | || |  :|| |:||
Db         21 NLDVEEPIVFREDAASFGQTVVQFGGSRLVVGAPLEAVAVNQTGRLYDCAPATGMCQPIV 80

Qy         62 LQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQKF 121
              |: |:|||||||||||   |:  |||||||  : | :| | || | | ||:| |  |
Db         81 LRSPLEAVNMSLGLSLVTATNNAQLLACGPTAQRACVKNMYAKGSCLLLGSSL-QFIQAV 139

Qy        122 PEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEFR 181
              | ::  ||::: |||||||||||   || :||:||  :| :   : |||||||||   :
Db        140 PASMPECPRQEMDIAFLIDGSGSINQRDFAQMKDFVKALMGEFASTSTLFSLMQYSNILK 199

Qy        182 IHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVIT 241
              |||| ||:|  :|:||| || || | |:||||:| |: || :  ||:||:| |||:|||
Db        200 THFTFTEFKNILDPQSLVDPIVQLQGLTYTATGIRTVMEELFHSKNGSRKSAKKILLVIT 259

Qy        242 DGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVNN 301
              ||:|: ||| | |||| ||: |:||| |||||||:   : :||||| | ||:||||:| |
Db        260 DGQKYRDPLEYSDVIPAADKAGIIRYAIGVGDAFQEPTALKELNTIGSAPPQDHVFKVGN 319

Qy        302 FEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAGG 361
              | ||::|| ||:||||||||: |||||:|||||||||:|:||:||:|  |||: |:||
Db        320 FAALRSIQRQLQEKIFAIEGTQSRSSSSFQHEMSQEGFSSALTSDGPVLGAVGSFSWSGG 379

Qy        362 VFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFRQ 421
              |||    : |||||:: : || |:||||: |:     | ||:||||||:|| | | :| |
Db        380 AFLYPPNTRPTFINMSQENVDMRDSYLGYSTAVAFWKGVHSLILGAPRHQHTGKVVIFTQ 439

Qy        422 NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPLP 481
                |   : |:|||||:|||||||||||| :|||||||||||||||||||||||| |:|
Db        440 EARHWRPKSEVRGTQIGSYFGASLCSVDVDRDGSTDLVLIGAPHYYEQTRGGQVSVFPVP 499

Qy        482 RGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFH 541
              | | ||||:| |:|||| |||||| |||||||||| | |||||||||::|||||:||
Db        500 -GVRGRWQCEATLHGEQGHPWGRFGVALTVLGDVNGDNLADVAIGAPGEEESRGAVYIFH 558

Qy        542 GTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQP 601
              | |   | || |||: ||:|| |||||||||||||| |||||| ||||||||||| |
Db        559 GASRLEIMPSPSQRVTGSQLSLRLQYFGQSLSGGQDLTQDGLVDLAVGAQGHVLLLRSLP 618
```

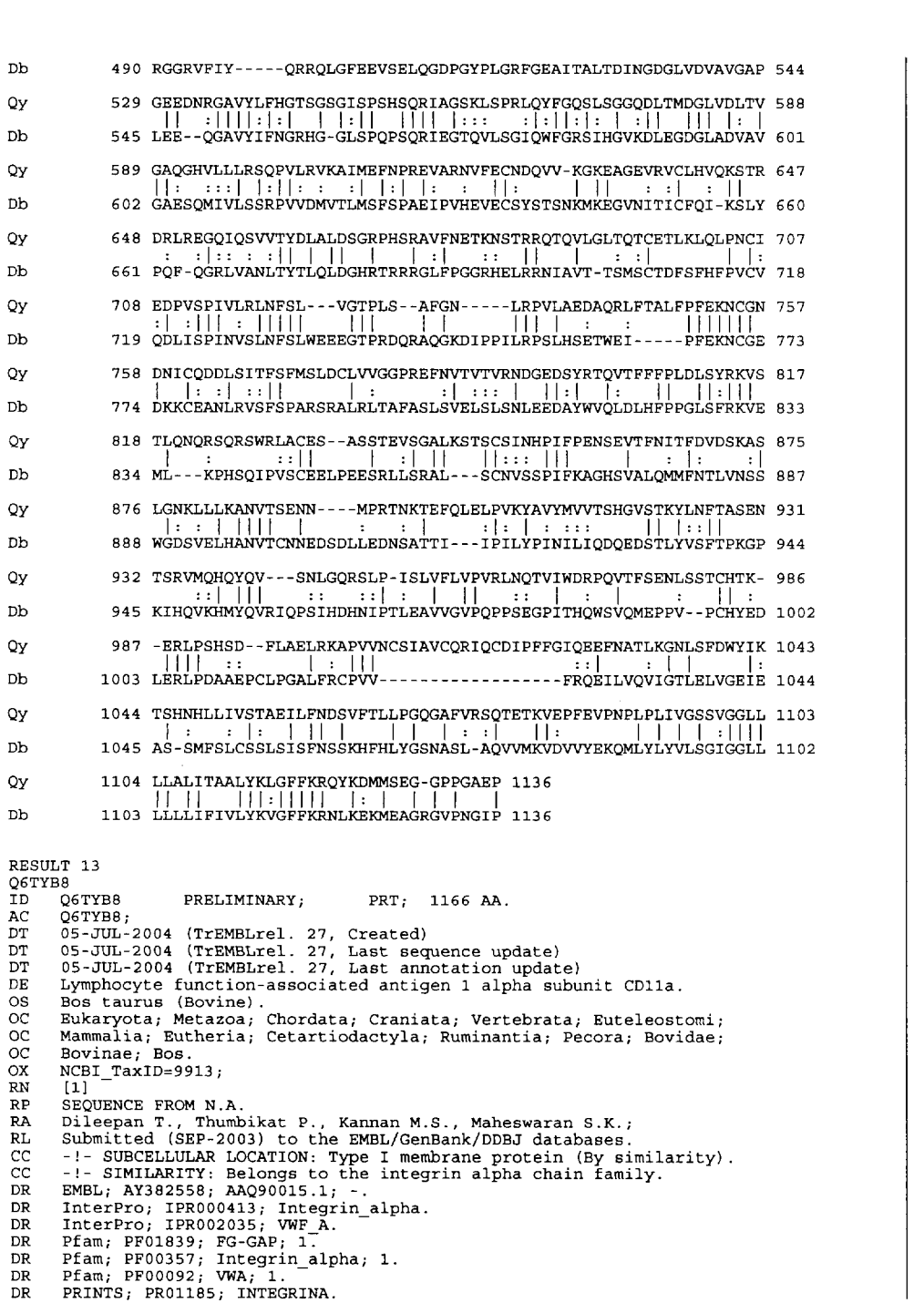
```
Qy        602 VLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVTY 661
              :|:|:  : | | |||: |::| ::     ||||  ||| | | : | |  | :|  | |
Db        619 LLKVELSIRFAPMEVAKAVYQCWERTPTVLEAGEATVCLTVHKGSPDLL--GNVQGSVRY 676

Qy        662 DLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNFS 721
              ||||| ||  |||:|:|||| |    : |||   |||:|| ||:|:|| ||||:||||||
Db        677 DLALDPGRLISRAIFDETKNCTLTGRKTLGLGDHCETVKLLLPDCVEDAVSPIILRLNFS 736

Qy        722 LVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVGG 781
              ||    |  || |||| :|   ||  |||||| : :|: || |:|:|  |  |||||
Db        737 LVRDSASP-RNLHPVLAVGSQDHITASLPFEKNCKQELLCEGDLGISFNFSGLQVLVVGG 795

Qy        782 PREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEVS 841
              |  ||||| |:||||| | | |::|  ||||:|:  | |  |  |||||:  : :
Db        796 SPELTVTVTVWNEGEDSYGTLVKFYYPAGLSYRRVTGTQ-QPHQYPLRLACEAEPAAQED 854

Qy        842 GALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEFQ 901
              |:|:|||||||| | :: || ||||  || ||::|||:|  :|||| | |||| ||
Db        855 --LRSSSCSINHPIFREGAKTTFMITFDVSYKAFLGDRLLLRAKASSENNKPDTNKTAFQ 912

Qy        902 LELPVKYAVYMVVTSHGVSTKYLNFTASENTSR-VMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||| || :::    || ::||::|    |    |:|:|:||    | : : | |||
Db        913 LELPVKYTVYTLISRQEDSTNHVNFSSSHGGRRQEAAHRYRVNNLSPLKLAVRVNFWVPV 972

Qy        961 RLNQTVIWDRPQVTFSENLSS-TCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              ||   :||   || |      :| :: : | : ||| ::::  |::|||| |   :|||
Db        973 LLNGVAVWD---VTLSSPAQGVSCVSQMKPPQNPDFLTQIQRRSVLDCSIADCLHFRCDI 1029

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              |   ||:| :  |:|||| |  :|    :|:|| ||| |: ||:: |||| ||:|:| |
Db       1030 PSLDIQDELDFILRGNLSFGWVSQTLQEKVLLVSEAEITFDTSVYSQLPGQEAFLRAQVE 1089

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMM 1126
              | :| : |  |: |: |||||||||||||  ||||||||||||:|:
Db       1090 TTLEEYVVYEPIFLVAGSSVGGLLLLALITVVLYKLGFFKRQYKEML 1136
```

```
RESULT 11
ITAL_BOVIN
ID   ITAL_BOVIN     STANDARD;      PRT;  1165 AA.
AC   P61625;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE   alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE   chain) (CD11a).
GN   Name=ITGAL;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=14697514; DOI=10.1016/j.gene.2003.09.043;
RA   Fett T., Zecchinon L., Baise E., Desmecht D.;
RT   "The bovine (Bos taurus) CD11a-encoding cDNA: molecular cloning,
RT   characterisation and comparison with the human and murine
RT   glycoproteins.";
RL   Gene 325:97-101(2004).
CC   -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC       ICAM3 and ICAM4. It is involved in a variety of immune phenomena
CC       including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC       mediated killing, and antibody dependent killing by granulocytes
CC       and monocytes (By similarity).
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC       associates with beta-2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY267467; AAP94035.1; -.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     23       Potential.
FT   CHAIN        24   1165       Integrin alpha-L.
FT   DOMAIN       24   1084       Extracellular (Potential).
FT   TRANSMEM   1085   1105       Potential.
FT   DOMAIN     1106   1165       Cytoplasmic (Potential).
FT   REPEAT       40     88       FG-GAP 1.
FT   REPEAT       89    146       FG-GAP 2.
FT   DOMAIN      153    324       VWFA.
FT   REPEAT      347    397       FG-GAP 3.
FT   REPEAT      398    452       FG-GAP 4.
FT   REPEAT      454    513       FG-GAP 5.
FT   REPEAT      515    572       FG-GAP 6.
FT   REPEAT      575    627       FG-GAP 7.
FT   CA_BIND     465    473       Potential.
FT   CA_BIND     527    535       Potential.
FT   CA_BIND     587    595       Potential.
FT   SITE       1111   1115       GFFKR motif.
FT   DISULFID     71     78       By similarity.
FT   DISULFID    108    126       By similarity.
FT   DISULFID    650    704       By similarity.
FT   DISULFID    768    774       By similarity.
FT   DISULFID    842    858       By similarity.
FT   DISULFID    994   1009       By similarity.
FT   DISULFID   1017   1048       By similarity.
FT   CARBOHYD     33     33       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     86     86       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    185    185       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    646    646       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    667    667       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    723    723       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    859    859       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    894    894       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    929    929       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1056   1056       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1067   1067       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   1165 AA;  128725 MW;  DAEB3A3F1E1463CB CRC64;
```

```
  Query Match             26.9%;  Score 1578;  DB 1;  Length 1165;
  Best Local Similarity   34.9%;  Pred. No. 2.9e-95;
  Matches  413;  Conservative  197;  Mismatches  466;  Indels  106;  Gaps   33;

Qy          1 FNLDTENAMTFQENARG--FGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58
              :||| :   |      |  ||  |:|: |: |||||| |    |  |:|||   || |
Db         24 YNLDVRHVQNFSPPLAGRHFGYRVLQV-GNGVVVGAPSE---GNSMGNLYQCQPETGDCL 79

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQP 118
              |: |     : ||::||   :    |||| | : :|| :| |: |||:|   |||
Db         80 PVTLS--SNYTSKYLGMTLATDPTSDNLLACDPGLSRTCDQNIYLSGLCYLIHENLRGPV 137

Qy        119 QKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSE 178
               :    : | : : |: || ||| |:   :| :: :|:  ||::|  |   |: :|:|
Db        138 LQGHPGYQECIKGNVDLVFLFDGSMSLQQDEFEKIVDFMKDVMKKLSNSSYQFAAVQFST 197

Qy        179 EFRIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILI 238
              ||  ||| ::    :| :|:  :  :   |:|   :  | :|:     ||| :| |:||
Db        198 YFRTEFTFLDYIRQKDPDALLAGVKHMRLLTNTFGAINYVAKEVFRPDLGARPDATKVLI 257
```

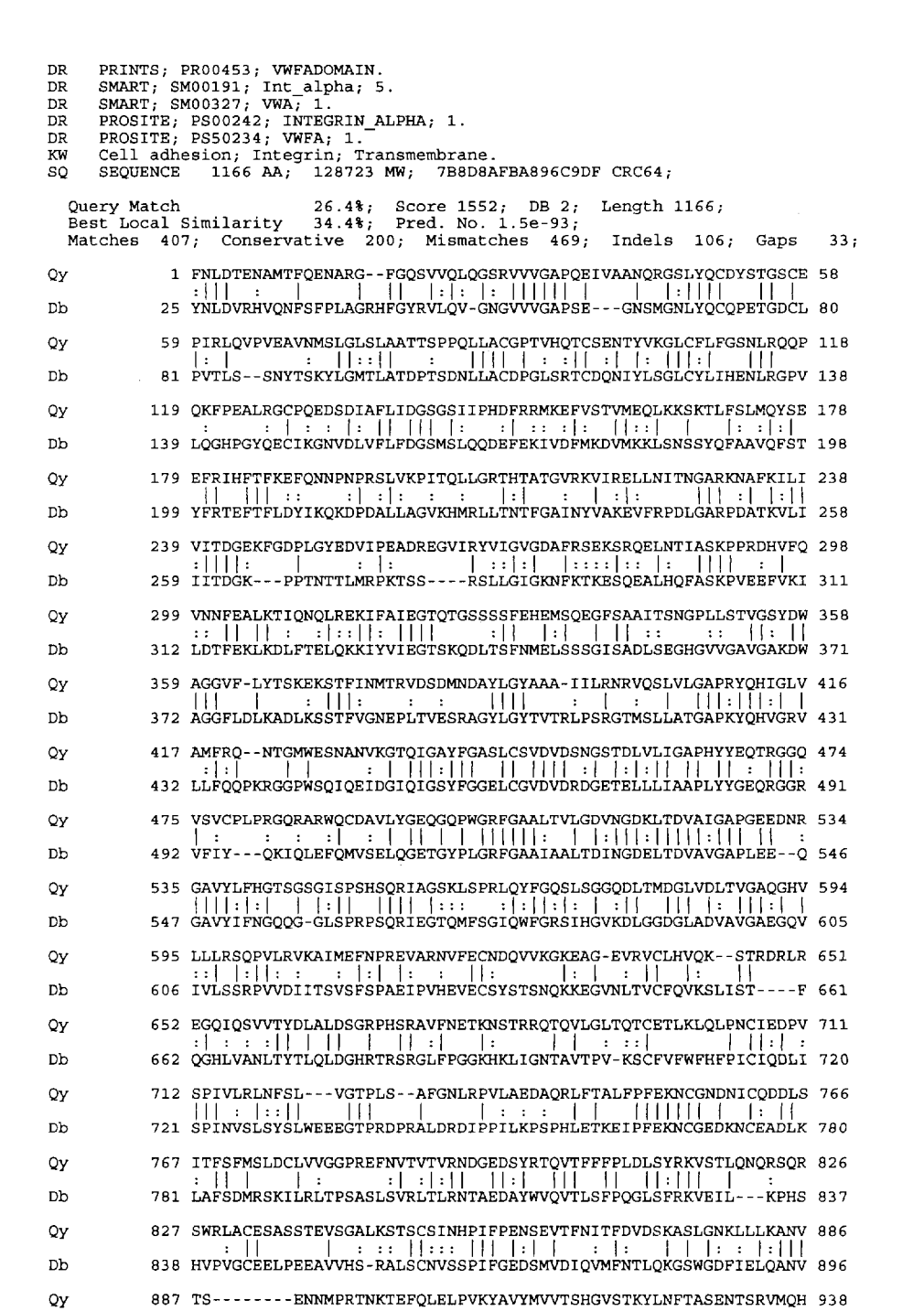
```
Qy        239 VITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQ 298
              :|||||         |  |  |  : :|||:||:|  |::::|:: |:  ||||  : |
Db        258 IITDGEATD-----EHNIDAA--KDIIRYIIGIGKNFKTKESQEALHQFASKPVEEFVKI 310

Qy        299 VNNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDW 358
              :: || || :  :|::||: ||||     :||  |:|  | || ::    ::  ||: ||
Db        311 LDTFEKLKDLFTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSEGHGVVGAVGAKDW 370

Qy        359 AGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQHIGLV 416
              |||   |    : |||:    :  :    ||||    :  |  :  |  |||||||:| |
Db        371 AGGFLDLKADLKSSTFVGNEPLTVESRAGYLGYTVTWLPSRGTMSLLATGAPRYQHVGRV 430

Qy        417 AMFRQ--NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQ 474
               :|:|    | |     : | |||:|||  || |||| :| |:|:|| || || : |||:
Db        431 LLFQQPKRGGPWSQIQEIDGIQIGSYFGGELCGVDVDRDGETELLLIAAPLYYGEQRGGR 490

Qy        475 VSVCPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNR 534
              | :    :  :  :|  : | || | | ||||||:  | |:|||:|||||:||| ||  :
Db        491 VFIY---QKIQLEFQMVSELQGETGYPLGRFGAAIAALTDINGDELTDVAVGAPLEE--Q 545

Qy        535 GAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHV 594
              ||||:|:|  | |:||  |||| |:::   :|:||:|: | :||  ||| |: |||:| |
Db        546 GAVYIFNGQQG-GLSPRPSQRIEGTQMFSGIQWFGRSIHGVKDLGGDGLADVAVGAEGQV 604

Qy        595 LLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAG-EVRVCLHVQK--STRDRLR 651
              ::| |:||: :   : |:| |:  :  ||:       |: |  : ||  |:   ||
Db        605 IVLSSRPVVDIITSVSFSPAEIPVHEVECSYSTSNQKKEGVNLTVCFQVKSLIST----F 660

Qy        652 EGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPV 711
              :| : : :|| | ||  |  || :|   |:      | |  : ::|       | ||:| :
Db        661 QGHLVANLTYTLQLDGHRTRSRGLFPGGKHKLIGNTAVTPV-KSCFVFWFHFPICIQDLI 719

Qy        712 SPIVLRLNFSL---VGTPLS--AFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLS 766
              ||| : |::||    |||     |      | : : :  | |   |||||||  |  |: ||
Db        720 SPINVSLSYSLWEEEGTPRDPRALDRDIPPILKPSPHLETKEIPFEKNCGEDKNCEADLK 779

Qy        767 ITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQR 826
              : || |    | :      :| :|:||  ||:|  |||  ||  ||:|||  |   :
Db        780 LAFSDMRSKILRLTPSASLSVRLTLRNTAEDAYWVQVTLSFPQGLSFRKVEIL---KPHS 836

Qy        827 SWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANV 886
                : ||      |  : :: ||::: ||| |:| |   : |:   | | |: : |:|||
Db        837 HVPVGCEELPEEAVVHS-RALSCNVSSPIFGEDSMVDIQVMFNTLQKGSWGDFIELQANV 895

Qy        887 TS--------ENNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQH 938
              :          |:|   |:       :|| | : ::      || |::||        ::|
Db        896 SCNNEDSSLLEDNSATTS-------IPVMYPINVLTKDQENSTLYISFTPKSPKIHHVKH 948

Qy        939 QYQV----SNLGQRSLPISLVFLVPVRLNQTVI---WD---RPQVTFS-ENLSSTCHTKE 987
               |||    ||      |:  :  ||   :: :|   |     | |  |  ||  |     |
Db        949 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPRNLESPSDEAE 1007

Qy        988 RLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHN 1047
                          :||       :| | |   ::|    : | :     || | :
Db       1008 ------------------SCSFGT--EFRCPIDF---RQEILVQVNGMVELRGTIKAS-S 1043

Qy       1048 HLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLAL 1107
               | : |:  | || |    | |: | : :|   ||:       | | | | :|||||  |
Db       1044 MLSLCSSLAISFNSSKHFHLYGRNASM-AQVVMKVDLVYEKEMLYLYVLSGIGGLLLLFL 1102

Qy       1108 ITAALYKLGFFKRQYKDMM-----------SEGGPPGAEPQ 1137
              |  ||||:|||||  |: |            : | |  | :
Db       1103 IFIALYKVGFFKRNLKEKMEANVDASSEIPGEDAGQPELEKE 1144


RESULT 12
ITAL_HUMAN
ID   ITAL_HUMAN     STANDARD;      PRT;  1170 AA.
AC   P20701; O43746;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE   alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE   chain) (CD11a).
GN   Name=ITGAL; Synonyms=CD11A;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RX   MEDLINE=89139587; PubMed=2537322;
RA   Larson R.S., Corbi A.L., Berman L., Springer T.;
RT   "Primary structure of the leukocyte function-associated molecule-1
RT   alpha subunit: an integrin with an embedded domain defining a protein
RT   superfamily.";
RL   J. Cell Biol. 108:703-712(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RX   MEDLINE=99425270; PubMed=10493829;
RA   Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
RA   Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA   Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
RA   Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT   "Genome duplications and other features in 12 Mb of DNA sequence from
RT   human chromosome 16p and 16q.";
RL   Genomics 60:295-308(1999).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.
RX   MEDLINE=96036067; PubMed=7479767;
RA   Qu A., Leahy D.J.;
RT   "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L
RT   beta 2) integrin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.
RX   MEDLINE=96398682; PubMed=8805579;
RA   Qu A., Leahy D.J.;
RT   "The role of the divalent cation in the structure of the I domain from
RT   the CD11a/CD18 integrin.";
RL   Structure 4:931-942(1996).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.
RX   MEDLINE=99425288; PubMed=10493852;
RA   Kallen J., Welzenbach K., Ramage P., Geyl D., Kriwacki R., Legge G.,
RA   Cottens S., Weitz-Schmidt G., Hommel U.;
RT   "Structural basis for LFA-1 inhibition upon lovastatin binding to the
RT   CD11a I-domain.";
RL   J. Mol. Biol. 292:1-9(1999).
CC   -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC       ICAM3 and ICAM4. It is involved in a variety of immune phenomena
CC       including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC       mediated killing, and antibody dependent killing by granulocytes
CC       and monocytes.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P20701-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P20701-2; Sequence=VSP_002738;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: Leukocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11a.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Y00796; CAA68747.1; -.
DR   EMBL; AC002310; AAC31672.1; -.
DR   PIR; S03308; S03308.
DR   PDB; 1CQP; X-ray; A/B=153-334.
DR   PDB; 1DGQ; NMR; A=149-336.
DR   PDB; 1LFA; X-ray; A/B=-.
DR   PDB; 1MJN; X-ray; A=153-331.
DR   PDB; 1MQ8; X-ray; B/D=155-331.
DR   PDB; 1MQ9; X-ray; A=152-330.
DR   PDB; 1MQA; X-ray; A=152-330.
DR   PDB; 1ZON; X-ray; @=150-336.
DR   PDB; 1ZOO; X-ray; A/B=150-336.
DR   PDB; 1ZOP; X-ray; A/B=150-336.
DR   Genew; HGNC:6148; ITGAL.
DR   MIM; 153370; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   3D-structure; Alternative splicing; Calcium; Cell adhesion;
KW   Direct protein sequencing; Glycoprotein; Integrin; Magnesium;
KW   Receptor; Repeat; Signal; Transmembrane.
FT   SIGNAL       1     25
FT   CHAIN       26   1170       Integrin alpha-L.
FT   DOMAIN      26   1090       Extracellular (Potential).
FT   TRANSMEM  1091   1111       Potential.
FT   DOMAIN    1112   1170       Cytoplasmic (Potential).
FT   REPEAT      42     91       FG-GAP 1.
FT   REPEAT      92    149       FG-GAP 2.
FT   DOMAIN     156    327       VWFA.
FT   REPEAT     350    400       FG-GAP 3.
FT   REPEAT     401    455       FG-GAP 4.
FT   REPEAT     457    516       FG-GAP 5.
FT   REPEAT     518    575       FG-GAP 6.
FT   REPEAT     578    630       FG-GAP 7.
FT   CA_BIND    468    476       Potential.
FT   CA_BIND    530    538       Potential.
FT   CA_BIND    590    598       Potential.
FT   SITE      1115   1119       GFFKR motif.
FT   DISULFID    73     80       By similarity.
FT   DISULFID   111    129       By similarity.
FT   DISULFID   653    707       By similarity.
FT   DISULFID   771    777       By similarity.
FT   DISULFID   845    861       By similarity.
FT   DISULFID   998   1013       By similarity.
FT   DISULFID  1021   1052       By similarity.
FT   CARBOHYD    65     65       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   188    188       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   649    649       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   670    670       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   726    726       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   730    730       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   862    862       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   885    885       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   897    897       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1060   1060       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1071   1071       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC   954    954       Q -> QGVHGLVEMQTSKQILCRPAGDAEHTVGAQEGELPC
FT                               PWGVSEAFRDNIRAGPCR (in isoform 2).
FT                               /FTId=VSP_002738.
FT   CONFLICT   214    214       R -> W (in Ref. 1 and 2).
FT   CONFLICT   660    660       Y -> I (in Ref. 2).
FT   STRAND     155    162
FT   STRAND     164    164
FT   TURN       165    166
FT   HELIX      169    185
FT   TURN       186    188
FT   STRAND     191    198
FT   STRAND     202    206
FT   HELIX      208    214
FT   HELIX      217    221
FT   TURN       222    223
FT   STRAND     229    229
FT   HELIX      233    243
FT   TURN       244    244
FT   HELIX      247    249
FT   TURN       250    250
FT   TURN       253    254
FT   STRAND     256    263
FT   HELIX      274    276
FT   TURN       277    278
FT   STRAND     280    286
FT   HELIX      288    290
FT   HELIX      293    297
FT   TURN       298    299
FT   HELIX      300    302
FT   HELIX      307    310
FT   STRAND     311    314
FT   HELIX      317    319
FT   HELIX      323    330
SQ   SEQUENCE   1170 AA;  128819 MW;  39A7AF92EF286FC0 CRC64;

  Query Match             26.4%;  Score 1552.5;  DB 1;  Length 1170;
  Best Local Similarity   34.5%;  Pred. No. 1.4e-93;
  Matches  405;  Conservative  208;  Mismatches  460;  Indels  101;  Gaps   37;

Qy          1 FNLDTENAMTFQ--ENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58
              :|||   | :|       | ||  |:|: |: |:|||| |    |  ||||||   || |
Db         26 YNLDVRGARSFSPPRAGRHFGYRVLQV-GNGVIVGAPGE---GNSTGSLYQCQSGTGHCL 81

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLR--- 115
              |: |:     : ||::||   :   :||| | : :|| :|||: |||:|| ||:
Db         82 PVTLR-GSNYTSKYLGMTLATDPTDGSILACDPGLSRTCDQNTYLSGLCYLFRQNLQGPM 140

Qy        116 -QQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLM 174
               |    | | ::|    : |: || ||| |: | :|::: :|:  ||::|  :   |: :
Db        141 LQGRPGFQECIKG----NVDLVFLFDGSMSLQPDEFQKILDFMKDVMKKLSNTSYQFAAV 196

Qy        175 QYSEEFRIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAF 234
              |:|  ::  | | ::    :| :|:| :  :|  |:|   :  | |:      ||| :|
Db        197 QFSTSYKTEFDFSDYVKRKDPDALLKHVKHMLLLTNTFGAINYVATEVFREELGARPDAT 256

Qy        235 KILIVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRD 294
              |:||:|||||       | |   :    :|||:||:|  |::::|:: |:  ||||  :
Db        257 KVLIIITDGE--ATDSGNIDAAKD-----IIRYIIGIGKHFQTKESQETLHKFASKPASE 309

Qy        295 HVFQVNNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVG 354
               |  :: || || :  :|::||: ||||     :||  |:|  | || ::    ::  ||
Db        310 FVKILDTFEKLKDLFTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSRGHAVVGAVG 369

Qy        355 SYDWAGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQH 412
              : ||||   |    :  |||    :  ::   ||||    :  | :   |  |||||||
Db        370 AKDWAGGFLDLKADLQDDTFIGNEPLTPEVRAGYLGYTVTWLPSRQKTSLLASGAPRYQH 429

Qy        413 IGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQT 470
              :| | :|:  |  | |     : |||||:|||  || |||| :| |:|:|||| :| :
Db        430 MGRVLLFQEPQGGGHWSQVQTIHGTQIGSYFGGELCGVDVDQDGETELLLIGAPLFYGEQ 489

Qy        471 RGGQVSVCPLPRGQRARWQCDAV--LYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAP 528
              |||:| :      || :   : |  | |: | | |||| |:| | |:||| | |||:|||
```



RESULT 13
Q6TYB8
ID Q6TYB8 PRELIMINARY; PRT; 1166 AA.
AC Q6TYB8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AY382558; AAQ90015.1; -.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 1.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Cell adhesion; Integrin; Transmembrane.
SQ SEQUENCE 1166 AA; 128723 MW; 7B8D8AFBA896C9DF CRC64;

Query Match 26.4%; Score 1552; DB 2; Length 1166;
Best Local Similarity 34.4%; Pred. No. 1.5e-93;
Matches 407; Conservative 200; Mismatches 469; Indels 106; Gaps 33;



```
Db        897 SCNNEDSSLLEDNSATTS-------IPVMYPINVLTKDQENSTLYISFTPKSPRIHHVKH 949

Qy        939 QYQV----SNLGQRSLPISLVFLVPVRLNQTVI---WD---RPQVTFS-ENLSSTCHTKE 987

Db        950 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPRNLESPSDEAE 1008

Qy        988 RLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHN 1047

Db       1009 ------------------SCSFGT--EFRCPIDF---RQEILVQVNGMVELRGTIKAS-S 1044

Qy       1048 HLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLAL 1107

Db       1045 MLSLCSSLAISFNSSKHFHLHGRNASM-AQVVMKVDLVYEKEMLYLYVLSGIGGLLLLFL 1103

Qy       1108 ITAALYKLGFFKRQYKDMM------------SEGGPPGAEPQ 1137

Db       1104 IFIALYKVGFFKRNLKEKMEANVDASSEIPGEDAGQPELEKE 1145


RESULT 14
AAQ90015
ID   AAQ90015    PRELIMINARY;      PRT;  1166 AA.
AC   AAQ90015;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT   "Molecular cloning and sequencing of bovine CD11a.";
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY382558; AAQ90015.1; -.
SQ   SEQUENCE   1166 AA;  128723 MW;  7B8D8AFBA896C9DF CRC64;

  Query Match             26.4%;  Score 1552;  DB 2;  Length 1166;
  Best Local Similarity   34.4%;  Pred. No. 1.5e-93;
  Matches  407;  Conservative  200;  Mismatches  469;  Indels  106;  Gaps   33;

Qy          1 FNLDTENAMTFQENARG--FGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCE 58

Db         25 YNLDVRHVQNFSFPLAGRHFGYRVLQV-GNGVVVGAPSE---GNSMGNLYQCQPETGDCL 80

Qy         59 PIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQP 118

Db         81 PVTLS--SNYTSKYLGMTLATDPTSDNLLACDPGLSRTCDQNIYLSGLCYLIHENLRGPV 138

Qy        119 QKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSE 178

Db        139 LQGHPGYQECIKGNVDLVFLFDGSMSLQQDEFEKIVDFMKDVMKKLSNSSYQFAAVQFST 198

Qy        179 EFRIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILI 238

Db        199 YFRTEFTFLDYIKQKDPDALLAGVKHMRLLTNTFGAINYVAKEVFRPDLGARPDATKVLI 258

Qy        239 VITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQ 298

Db        259 IITDGK---PPTNTTLMRPKTSS----RSLLGIGKNFKTKESQEALHQFASKPVEEFVKI 311

Qy        299 VNNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDW 358

Db        312 LDTFEKLKDLFTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSEGHGVVGAVGAKDW 371

Qy        359 AGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPRYQHIGLV 416

Db        372 AGGFLDLKADLKSSTFVGNEPLTVESRAGYLGYTVTRLPSRGTMSLLATGAPKYQHVGRV 431

Qy        417 AMFRQ--NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQ 474

Db        432 LLFQQPKRGGPWSQIQEIDGIQIGSYFGGELCGVDVDRDGETELLLIAAPLYYGEQRGGR 491

Qy        475 VSVCPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNR 534

Db        492 VFIY---QKIQLEFQMVSELQGETGYPLGRFGAAIAALTDINGDELTDVAVGAPLEE--Q 546

Qy        535 GAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHV 594

Db        547 GAVYIFNGQQG-GLSPRPSQRIEGTQMFSGIQWFGRSIHGVKDLGGDGLADVAVGAEGQV 605

Qy        595 LLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAG-EVRVCLHVQK--STRDRLR 651

Db        606 IVLSSRPVVDIITSVSFSPAEIPVHEVECSYSTSNQKKEGVNLTVCFQVKSLIST----F 661

Qy        652 EGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPV 711

Db        662 QGHLVANLTYTLQLDGHRTRSRGLFPGGKHKLIGNTAVTPV-KSCFVFWFHFPICIQDLI 720

Qy        712 SPIVLRLNFSL---VGTPLS--AFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLS 766

Db        721 SPINVSLSYSLWEEEGTPRDPRALDRDIPPILKPSPHLETKEIPFEKNCGEDKNCEADLK 780

Qy        767 ITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQR 826

Db        781 LAFSDMRSKILRLTPSASLSVRLTLRNTAEDAYWVQVTLSFPQGLSFRKVEIL---KPHS 837

Qy        827 SWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANV 886

Db        838 HVPVGCEELPEEAVVHS-RALSCNVSSPIFGEDSMVDIQVMFNTLQKGSWGDFIELQANV 896

Qy        887 TS--------ENNMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQH 938

Db        897 SCNNEDSSLLEDNSATTS-------IPVMYPINVLTKDQENSTLYISFTPKSPRIHHVKH 949

Qy        939 QYQV----SNLGQRSLPISLVFLVPVRLNQTVI---WD---RPQVTFS-ENLSSTCHTKE 987

Db        950 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPRNLESPSDEAE 1008

Qy        988 RLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHN 1047

Db       1009 ------------------SCSFGT--EFRCPIDF---RQEILVQVNGMVELRGTIKAS-S 1044

Qy       1048 HLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLLAL 1107

Db       1045 MLSLCSSLAISFNSSKHFHLHGRNASM-AQVVMKVDLVYEKEMLYLYVLSGIGGLLLLFL 1103

Qy       1108 ITAALYKLGFFKRQYKDMM------------SEGGPPGAEPQ 1137

Db       1104 IFIALYKVGFFKRNLKEKMEANVDASSEIPGEDAGQPELEKE 1145


RESULT 15
ITAL_MOUSE
ID   ITAL_MOUSE     STANDARD;      PRT;  1163 AA.
AC   P24063;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE   alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE   chain) (CD11a).
GN   Name=Itgal; Synonyms=Lfa-1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91268576; PubMed=2051027;
RA   Kaufmann Y., Tseng E., Springer T.A.;
RT   "Cloning of the murine lymphocyte function-associated molecule-1
```

```
RT   alpha-subunit and its expression in COS cells.";
RL   J. Immunol. 147:369-374(1991).
RN   [2]
RP   SEQUENCE OF 24-42.
RX   MEDLINE=85188276; PubMed=3887182;
RA   Springer T.A., Teplow D.B., Dreyer W.J.;
RT   "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
RT   glycoproteins and unexpected relation to leukocyte interferon.";
RL   Nature 314:540-542(1985).
CC   -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC       ICAM3 and ICAM4. Is involved in a variety of immune phenomena
CC       including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC       mediated killing, and antibody dependent killing by granulocytes
CC       and monocytes. Mice expressing a null mutation of the alpha-L
CC       subunit gene demonstrate impaired tumor rejection and impaired
CC       leukocytes recruitment.
CC   -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC       associates with beta-2.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Leukocytes.
CC   -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC       with I-domains do not undergo protease cleavage.
CC   -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC   -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M60778; AAA39426.1; -.
DR   PIR; I56126; I56126.
DR   HSSP; P20701; 1DGQ.
DR   MGD; MGI:96606; Itgal.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 2.
DR   Pfam; PF00357; Integrin_alpha; 1.
DR   Pfam; PF00092; VWA; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 5.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Calcium; Cell adhesion; Direct protein sequencing; Glycoprotein;
KW   Integrin; Magnesium; Receptor; Repeat; Signal; Transmembrane.
FT   SIGNAL       1     23
FT   CHAIN       24   1163       Integrin alpha-L.
FT   DOMAIN      24   1087       Extracellular (Potential).
FT   TRANSMEM  1088   1108       Potential.
FT   DOMAIN    1109   1163       Cytoplasmic (Potential).
FT   REPEAT      39     88       FG-GAP 1.
FT   REPEAT      89    146       FG-GAP 2.
FT   DOMAIN     153    325       VWFA.
FT   REPEAT     348    398       FG-GAP 3.
FT   REPEAT     399    454       FG-GAP 4.
FT   REPEAT     455    514       FG-GAP 5.
FT   REPEAT     516    573       FG-GAP 6.
FT   REPEAT     576    628       FG-GAP 7.
FT   CA_BIND    466    474       Potential.
FT   CA_BIND    528    536       Potential.
FT   CA_BIND    588    596       Potential.
FT   SITE      1111   1115       GFFKR motif.
FT   DISULFID    70     77       By similarity.
FT   DISULFID   108    126       By similarity.
FT   DISULFID   147    199       By similarity.
FT   DISULFID   651    705       By similarity.
FT   DISULFID   767    773       By similarity.
FT   DISULFID   840    856       By similarity.
FT   DISULFID   993   1009       By similarity.
FT   DISULFID  1017   1048       By similarity.
FT   CARBOHYD    86     86       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   185    185       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   270    270       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   444    444       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   668    668       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   696    696       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   724    724       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   728    728       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   776    776       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   857    857       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   880    880       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   890    890       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   899    899       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   927    927       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1056   1056       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE  1163 AA;  128343 MW;  A7A3078489E8232F CRC64;

  Query Match             26.1%;  Score 1533.5;  DB 1;  Length 1163;
  Best Local Similarity   34.0%;  Pred. No. 2.5e-92;
  Matches  398;  Conservative  217;  Mismatches  458;  Indels   99;  Gaps   36;

Qy          1 FNLDTENAMTFQENA-RGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEP 59
              :||||     :|   | | ||  |:|::   ||||||  |     :  | || |  |:  |:|
Db         24 YNLDTRPTQSFLAQAGRHFGYQVLQIEDG-VVVGAPGE---GDNTGGLYHCRTSSEFCQP 79

Qy         60 IRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQ 119
              : |      :  ||::||   :   |||| | : :|| :|||: |||:||  :|
Db         80 VSLH-GSNHTSKYLGMTLATDAAKGSLLACDPGLSRTCDQNTYLSGLCYLFPQSLEGPML 138

Qy        120 KFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEE 179
              :   | : | :   |: || ||| |:   || :: ||:  || :|  :   |: :|:| :
Db        139 QNRPAYQECMKGKVDLVFLFDGSQSLDRKDFEKILEFMKDVMRKLSNTSYQFAAVQFSTD 198

Qy        180 FRIHFTFKEF-QNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILI 238
              |  ||| :: : | ||  |:  :  :   |:|   :  |:  :    :||| :| |:|:
Db        199 CRTEFTFLDYVKQNKNPDVLLGSVQPMFLLTNTFRAINYVVAHVFKEESGARPDATKVLV 258

Qy        239 VITDGEKFGDPLGYEDVIPEADREG-------VIRYVIGVGDAFRSEKSRQELNTIASKP 291
              :||||              ||  :|       : ||:||:|  | | : :: |:  ||:|
Db        259 IITDG--------------EASDKGNISAAHDITRYIIGIGKHFVSVQKQKTLHIFASEP 304

Qy        292 PRDHVFQVNNFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLS 351
              : |  :: || || :   |: :|:|||||      :||  |:|  | || ::    ::
Db        305 VEEFVKILDTFEKLKDLFTDLQRRIYAIEGTNRQDLTSFNMELSSSGISADLSKGHAVVG 364

Qy        352 TVGSYDWAGGVF-LYTSKEKSTFINMTRVDSDMNDAYLGYAAA-IILRNRVQSLVLGAPR 409
              ||: |||||    |    : :||:    : ||:   ||||  | :  |:     |  ||||
Db        365 AVGAKDWAGGFLDLREDLQGATFVGQEPLTSDVRGGYLGYTVAWMTSRSSRPLLAAGAPR 424

Qy        410 YQHIGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYY 467
              |||:|  | :|:  :  | |     ::|||||:|||  |||||:| :|  :|:|||| ::
Db        425 YQHVGQVLLFQAPEAGGRWNQTQKIEGTQIGSYFGGELCSVDLDQDGEAELLLIGAPLFF 484

Qy        468 EQTRGGQVSVCPLPRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGA 527
              : |||:|      : ::: ::  : | |: | | ||||||:| | |:|||:|||||:||
Db        485 GEQRGGRVFTY---QRRQSLFEMVSELQGDPGYPLGRFGAAITALTDINGDRLTDVAVGA 541

Qy        528 PGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLT 587
              | ||  :||||:|:|  | |:||  |||| |::: | :::||:|: | :||  | | |:
Db        542 PLEE--QGAVYIFNGKPG-GLSPQPSQRIQGAQVFPGIRWFGRSIHGVKDLGGDRLADVV 598

Qy        588 VGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGKEAG-EVRVCLHVQKST 646
              |||:| |::| |:||:  |   : |:| |:  :  ||:    : :: | ::: |  ::  |
Db        599 VGAEGRVVVLSSRPVVDVVTELSFSPEEIPVHEVECSYSAREEQKHGVKLKACFRIKPLT 658

Qy        647 RDRLREGQIQSVVTYDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNC 706
                :|:: : ::| | ||  |  || :| :  :     | :   ::|       | |
Db        659 PQ--FQGRLLANLSYTLQLDGHRMRSRGLFPDGSHELSGNTSITP-DKSCLDFHFHFPIC 715

Qy        707 IEDPVSPIVLRLNFSLV---GTPLSAFGN-LRPVLAEDAQRLFTALFPFEKNCGNDNICQ 762
```

```
Db        716 IQDLISPINVSLNFSLLEEEGTPRDQKGRAMQPILRPSIHTV-TKEIPFEKNCGEDKKCE 774

Qy        763 DDLSITFSFMSLDCLVVGGP------REFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKV 816
Db        775 ANLTLSSPARS-------GPLRLMSSASLAVEWTLSNSGEDAYWVRLDLDFPRGLSFRKV 827

Qy        817 STLQNQRSQRSWRLACESASSTEVSGAL-KSTSCSINHPIFPENSEVTFNITFDVDSKAS 875
Db        828 EMLQ---PHSRMPVSCEEL--TEGSSLLTKTLKCNVSSPIFKAGQEVSLQVMFNTLLNSS 882

Qy        876 LGNKLLLKANVTSEN-NMPRTNKTEFQLELPVKYAVYMVVTSHGVSTKYLNFTASENTSR 934
Db        883 WEDFVELNGTVHCENENSSLQEDNSAATHIPVLYPVNILTKEQENSTLYISFTPKGPKTQ 942

Qy        935 VMQHQYQVSNLGQRSLPISLVFLVPVRLNQTVIWDRPQ----VTFSENLSS----TCHTK 986
Db        943 QVQHVYQV-----RIQPSAYDHNMPT-LEALVGVPRPHSEDLITYTWSVQTDPLVTCHSE 996

Qy        987 E-RLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTS 1045
Db        997 DLKRPSSE---AEQPCLPGV--------QFRCPIVF---RWEILIQVTGTVELSKEIKAS 1042

Qy       1046 HNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTETKVEPFEVPNPLPLIVGSSVGGLLLL 1105
Db       1043 -STLSLCSSLSVSFNSSKHFHLYGSKA-SEAQVLVKVDLIHEKEMLHVYVLSGIGGLVLL 1100

Qy       1106 ALITAALYKLGFFKRQYKDMM-SEGGPPGAEP 1136
Db       1101 FLIFLALYKVGFFKRNLKEKMEADGGVPNGSP 1132
```

Search completed: November 9, 2004, 12:17:43
Job time : 141.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:45:42 ; Search time 103.25 Seconds
(without alignments)
3950.365 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5868	99.9	1153	2	AAW65090	Aaw65090 Human Bet
2	5868	99.9	1153	3	AAB07360	Aab07360 Human CD1
3	5868	99.9	1153	5	AAU80252	Aau80252 Human int
4	5868	99.9	1153	5	ABG61469	Abg61469 Human Bet
5	5868	99.9	1153	5	AAO14428	Aao14428 Integrin
6	5868	99.9	1153	7	ADD25615	Add25615 Binding d
7	5858	99.7	1153	2	AAR04136	Aar04136 Alpha sub
8	5852.5	99.6	1152	8	ADM99589	Adm99589 Human int
9	5845.5	99.5	1152	8	ADP12435	Adp12435 Protein e
10	3487	59.4	1163	8	ADP44061	Adp44061 Human CD1
11	3473	59.1	1163	8	ADN02004	Adn02004 Human inf
12	3473	59.1	1163	8	ADQ17510	Adq17510 Human sof
13	3464	59.0	1163	2	AAR07120	Aar07120 p150.95 a
14	3450	58.7	1163	2	AAW65091	Aaw65091 Human Bet
15	3450	58.7	1163	3	AAB07361	Aab07361 Human CD1
16	3450	58.7	1163	5	ABG61470	Abg61470 Human Bet
17	3448	58.7	1163	6	ABU07406	Abu07406 Protein d
18	3448	58.7	1163	7	ADG32005	Adg32005 Human hom
19	3417	58.2	1161	2	AAR78166	Aar78166 Human bet
20	3417	58.2	1161	2	AAW23049	Aaw23049 Human bet
21	3417	58.2	1161	2	AAW57491	Aaw57491 Human bet
22	3417	58.2	1161	2	AAW65089	Aaw65089 Human Bet
23	3417	58.2	1161	2	AAW72825	Aaw72825 Human alp
24	3417	58.2	1161	2	AAW73342	Aaw73342 Human alp
25	3417	58.2	1161	3	AAB07359	Aab07359 Human alp
26	3417	58.2	1161	5	ABG61468	Abg61468 Human Bet
27	3401.5	57.9	1161	2	AAW23064	Aaw23064 Human bet
28	3401.5	57.9	1161	2	AAW65106	Aaw65106 Human Bet
29	3401.5	57.9	1161	2	AAW72837	Aaw72837 Human alp
30	3401.5	57.9	1161	2	AAW73343	Aaw73343 Human alp
31	3401.5	57.9	1161	3	AAB07376	Aab07376 Human alp
32	3401.5	57.9	1161	5	ABG61485	Abg61485 Human Bet
33	3241.5	55.2	1161	2	AAR78169	Aar78169 Rat alpha
34	3239.5	55.1	1161	2	AAW23062	Aaw23062 Rat beta
35	3239.5	55.1	1161	2	AAW60004	Aaw60004 Rat alpha
36	3239.5	55.1	1161	2	AAW72824	Aaw72824 Rat alpha
37	3239.5	55.1	1161	3	AAB07374	Aab07374 Rat alpha
38	3239.5	55.1	1161	5	ABG61483	Abg61483 Rat Beta2
39	3232.5	55.0	1161	2	AAW65104	Aaw65104 Rat beta-
40	3232.5	55.0	1161	2	AAW73345	Aaw73345 Rat alpha
41	3228	54.9	1161	2	AAW23061	Aaw23061 Mouse bet
42	3228	54.9	1161	2	AAW60003	Aaw60003 Mouse alp
43	3228	54.9	1161	2	AAW65103	Aaw65103 Mouse bet
44	3228	54.9	1161	2	AAW72836	Aaw72836 Mouse alp
45	3228	54.9	1161	2	AAW73347	Aaw73347 Mouse alp

ALIGNMENTS

RESULT 1
AAW65090
ID AAW65090 standard; protein; 1153 AA.
XX
AC AAW65090;
XX
DT 28-SEP-1998 (first entry)
XX
DE Human Beta-integrin CD11b subunit protein.
XX
KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN US5728533-A.
XX
PD 17-MAR-1998.
XX
PF 07-JUN-1995; 95US-00485618.
XX
PR 23-DEC-1993; 93US-00173497.
PR 05-AUG-1994; 94US-00286889.
PR 21-DEC-1994; 94US-00362652.
XX
PA (ICOS-) ICOS CORP.
XX
PI Van Der Vieren M, Gallatin WM;
XX
DR WPI; 1998-206565/18.
XX
PT Screening assay for modulators of integrin binding - using immobilised or
PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.
XX
PS Example 5; Fig 1A-D; 106pp; English.
XX
CC This sequence represents a human beta-integrin CD11b subunit which is
CC used to describe a method for identifying compounds that modulate the
CC interaction of the beta-integrin alpha-d subunit with a binding partner
CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha
CC -d binding partner, one of which is immobilised and the other of which is
CC labelled, in the presence of a test compound, and determining if the
CC compound affects binding between the alpha-d polypeptide and alpha-d
CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment
CC comprising the cytoplasmic, transmembrane or extracellular domain of
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

```
CC   diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC   asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC   and rheumatoid arthritis
XX
SQ   Sequence 1153 AA;

  Query Match             99.9%;  Score 5868;  DB 2;  Length 1153;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              ||||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 2
AAB07360
ID   AAB07360 standard; protein; 1153 AA.
XX
AC   AAB07360;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Human CD11b protein sequence.
XX
KW   Human; macrophage infiltration inhibition; alpha_d integrin;
KW   leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW   inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW   atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW   lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW   rheumatoid arthritis; central nervous system injury; CD11b.
XX
OS   Homo sapiens.
XX
PN   WO200029446-A1.
XX
PD   25-MAY-2000.
XX
PF   16-NOV-1999;   99WO-US027139.
XX
PR   16-NOV-1998;   98US-00193043.
PR   08-JUL-1999;   99US-00350259.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin MW,  Van Der Vieren M;
XX
DR   WPI; 2000-387751/33.
XX
PT   Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT   macrophage infiltration and reduce inflammation at central nervous system
PT   injury sites.
XX
PS   Example 5; Fig 1; 270pp; English.
XX
CC   Integrins are a class of membrane-associated molecules that participate
CC   in cellular adhesion. Integrins are made up of an alpha subunit and a
CC   beta subunit. One class of human integrins are restricted to expression
CC   in white blood cells and have a common beta2 subunit: the leukocyte
CC   integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC   have an important role in immune and inflammatory responses. The present
CC   protein sequence is the human integrin alpha subunit CD11b. This sequence
CC   was used in an alignment to identify a novel beta2 integrin alpha
CC   subunit: alpha_d (AAA60014 and AAB07359). The present sequence has
CC   approximately 60% identity to the protein sequence of alpha_d. The
CC   Alpha_d gene and protein may be useful in therapy for diseases linked to
CC   aberrant alpha_d function e.g. Type I diabetes, atherosclerosis, multiple
CC   sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
CC   distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
CC   (LAD). In addition, anti-alpha_d monoclonal antibodies may be used in the
CC   inhibition of macrophage infiltration at the site of a central nervous
CC   system injury. The monoclonal antibodies can also be used to detect and
CC   diagnose Crohn's disease
XX
```

```
SQ   Sequence 1153 AA;

  Query Match            99.9%;  Score 5868;  DB 3;  Length 1153;
  Best Local Similarity  99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              ||||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 3
AAU80252
ID AAU80252 standard; protein; 1153 AA.
XX
AC AAU80252;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human integrin 1 alpha-M subunit protein.
XX
KW Integrin; antiinflammatory; imunosuppression; nephritis; dermatitis;
KW inflammatory disease; autoimmune disorder; Crohn's disease;
KW human immunodeficiency virus; HIV; myocardial infarction;
KW Sjorgen's syndrome; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT Misc-difference 499. .500
FT   /note= "Encoded by GGG CAG AGG"
XX
PN WO200218583-A2.
XX
PD 07-MAR-2002.
XX
PF 31-AUG-2001; 2001WO-US027227.
XX
PR 01-SEP-2000; 2000US-0229700P.
XX
PA (BLOO-) CENT BLOOD RES INC.
XX
PI Springer TA, Shimoaka M, Lu C;
XX
DR WPI; 2002-382964/41.
DR N-PSDB; ABK50046.
XX
PT Modified integrin-I or integrin I-like domain polypeptide useful as an
PT immunogen to produce antibodies specific to polypeptide, comprises a
PT disulfide bond such that polypeptide is stabilized in a desired
PT conformation.
XX
PS Disclosure; Page 109-112; 112pp; English.
XX
CC This invention relates to a modified integrin-I or integrin I-like domain
CC polypeptide comprising at least one disulfide bond so that the domain is
CC stabilised in a desired conformation. The polypeptide of the invention
CC may have antiinflammatory or imunosuppressive activities. The
CC polypeptides of the invention have an open conformation and are useful as
CC immunogens to produce antibodies that selectively bind to integrin I-
CC domain; and for identifying a modulator of integrin activity, or of
CC interaction of an integrin and a cognate ligand. The polypeptide of the
CC invention, or antibodies (preferably anti-LFA-1 antibody) is useful for
CC treating or preventing an integrin mediated disorder which is an
CC inflammatory or autoimmune disorder in a subject and for inhibiting the
CC binding of an integrin to a cognate ligand such as Crohn's disease,
CC nephritis; human immunodeficiency virus (HIV), myocardial infarction,
CC Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic
CC composition comprising the peptide of the invention is useful for
CC treating an integrin mediated disorder in a subject. The polypeptides
CC and/or active or antigenic fragments are useful as reagents for diagnosis
CC of integrin-mediated disorders. The present sequence represents the human
CC integrin-1 alpha-M protein subunit used to generate the mutant
CC polypeptides of the invention

```
XX
SQ   Sequence 1153 AA;

  Query Match           99.9%;  Score 5868;  DB 5;  Length 1153;
  Best Local Similarity 99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |||||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 4
ABG61469
ID   ABG61469 standard; protein; 1153 AA.
XX
AC   ABG61469;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human Beta2 integrin alphaCD11b subunit.
XX
KW   Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; LAD;
KW   leukocyte adhesion deficiency; inflammatory response; diabetes;
KW   multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW   inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW   immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
KW   intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW   locomotor recovery; locomotor damage; locomotor impairment;
KW   autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX
OS   Homo sapiens.
XX
PN   WO200230980-A2.
XX
PD   18-APR-2002.
XX
PF   15-OCT-2001; 2001WO-US032059.
XX
PR   13-OCT-2000; 2000US-00688307.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin WM,  Van Der Vieren M;
XX
DR   WPI; 2002-463260/49.
XX
PT   Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT   recovery, inhibiting locomotor damage, limiting locomotor impairment, or
PT   limiting autonomic and sensory dysfunction following spinal cord injury.
XX
PS   Example 5; Page 191-194; 270pp; English.
XX
CC   The invention relates to promoting locomotor recovery, inhibiting
CC   locomotor damage, limiting locomotor impairment, or limiting autonomic
CC   and sensory dysfunction following spinal cord injury by administering an
CC   anti-alpha_d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
CC   spinal cord injury victim. The method also involves the use of a ligand
CC   selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,
CC   vascular cell adhesion molecule). The method is useful for promoting
CC   locomotor recovery, inhibiting locomotor damage, limiting locomotor
CC   impairment, or limiting autonomic and sensory dysfunction following
CC   spinal cord injury. In particular, the spinal cord injury comprises
CC   compression of the spinal cord. The antibodies are also useful for
CC   reducing inflammation at the site of a central nervous system injury. The
CC   specification also details the identification of Beta2 integrin alphaD
CC   cDNAs and proteins, for use in raising the antibodies. Beta2 integrins
CC   are implicated in diseases such as LAD (leukocyte adhesion deficiency,
CC   inflammatory response, diabetes, multiple sclerosis, arthritis, graft
CC   atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative
CC   colitis, immune complex alveolitis and leukaemia. The present sequence is
CC   a Beta2 integrin alpha subunit sequence included for comparison with the
CC   Beta2 integrin alphaD protein sequences
XX
```

```
SQ   Sequence 1153 AA;

  Query Match             99.9%;  Score 5868;  DB 5;  Length 1153;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              ||||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 5
AAO14428
ID   AAO14428 standard; protein; 1153 AA.
XX
AC   AAO14428;
XX
DT   03-MAY-2002  (first entry)
XX
DE   Integrin Mac-1 alpha subunit.
XX
KW   Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
KW   open conformation; integrin related inflammatory disorder;
KW   integrin related immunological disorder; rheumatoid arthritis; ischaemia;
KW   reperfusion; hypovolemic shock; infarction; cerebral shock;
KW   viral infection; cancer; gene therapy; vaccine;
KW   bioactive agent screening.
XX
OS   Unidentified.
XX
PN   WO200204521-A2.
XX
PD   17-JAN-2002.
XX
PF   09-JUL-2001; 2001WO-US021805.
XX
PR   07-JUL-2000; 2000US-0216600P.
XX
PA   (CALY ) CALIFORNIA INST OF TECHNOLOGY.
PA   (BLOO-) CENT BLOOD RES.
XX
PI   Springer T;
XX
DR   WPI; 2002-148167/19.
XX
PT   New integrin I domain protein having alteration in at least 2
PT   noncontiguous regions and exits in an open conformation, useful for
PT   treating, preventing or suppressing inflammatory or immunological
PT   disorders.
XX
PS   Example 1; Fig 1F; 90pp; English.
XX
CC   The invention comprises structurally biased variant integrin inserted (I)
CC   domain proteins, wherein the alterations to the protein occur in at least
CC   two noncontinuous regions. Specifically the the variant integrin I domain
CC   proteins are structurally biased to exist in the open conformation,
CC   thereby altering the binding ability of the protein. The invention also
CC   comprises nucleic acids encoding the variant integrin I domain proteins.
CC   The integrin I domain proteins and nucleic acids are useful for treating,
CC   preventing or suppressing integrin related inflammatory and immunological
CC   disorders (e.g. rheumatoid arthritis). The variant integrin I domain
CC   proteins and nucleic acids can also be used for treating: ischaemia/
CC   reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral
CC   infection; and cancer. The variant integrin I domain nucleic acids and
CC   proteins may be used in gene therapy, as vaccines and to screen for
CC   bioactive agents. The present amino acid sequence represents the Mac-1
CC   alpha subunit of integrin
XX
SQ   Sequence 1153 AA;

  Query Match             99.9%;  Score 5868;  DB 5;  Length 1153;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              ||||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX
AC ADD25615;
XX
DT 15-JAN-2004 (first entry)
XX
DE Binding domain-immunoglobulin fusion protein-associated protein #85.
XX
KW Binding domain; immunoglobulin; fusion protein; cytostatic;
KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;
KW neuroprotective; hinge region; immunoglobulin heavy chain;
KW CH2 constant region; CH3 constant region; IgG1;
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.
XX
OS Unidentified.
XX
PN US2003118592-A1.
XX
PD 26-JUN-2003.
XX
PF 25-JUL-2002; 2002US-00207655.
XX
PR 17-JAN-2001; 2001US-0367358P.
PR 17-JAN-2002; 2002US-00053530.
PR 03-JUN-2002; 2002US-0385691P.
XX
PA (GENE-) GENECRAFT INC.
XX
PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
XX
DR WPI; 2003-801317/75.
XX
PT New binding domain-immunoglobulin fusion protein, useful for treating a
PT subject having or suspected of having a malignant condition or a B-cell
PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX
PS Disclosure; SEQ ID NO 176; 157pp; English.
XX
CC The invention relates to a binding domain-immunoglobulin fusion protein
CC comprising a binding domain polypeptide that is fused to an
CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
CC CH2 constant region polypeptide that is fused to the hinge region
CC polypeptide, and an immunoglobulin heavy chain CH3 constant region
CC polypeptide that is fused to the CH2 constant region polypeptide. The
CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin
CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge
CC region polypeptide, derived from (a) having 3 or more cysteine residues;
CC where the mutated human IgG1 immunoglobulin hinge region polypeptide
CC contains 2 cysteine residues, where the first cysteine is not mutated; a
CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
CC (a) having 3 or more cysteine residues, where the mutated human IgG1
CC immunoglobulin hinge region polypeptide contains no more than one
CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
CC polypeptide, derived from (a) having 3 or more cysteine residues; where
CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains
CC no cysteine residues. The binding domain-immunoglobulin fusion protein is
CC capable of at least one immunological activity comprising antibody
CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The
CC binding domain polypeptide is capable of specifically binding to an
CC antigen. Also included are an isolated polynucleotide encoding the
CC binding domain-immunoglobulin fusion protein, a recombinant expression
CC construct comprising the polynucleotide (operably linked to a promoter),
CC a host cell transformed or transfected with a recombinant expression
CC construct, producing the binding domain-immunoglobulin fusion protein, a

```
CC   pharmaceutical composition comprising the binding domain-immunoglobulin
CC   fusion protein or polynucleotide and a carrier, and treating a subject
CC   having or suspected of having a malignant condition or a B-cell disorder.
CC   The binding domain-immunoglobulin fusion protein is useful for treating a
CC   subject having or suspected of having a malignant condition or a B-cell
CC   disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,
CC   myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple
CC   sclerosis or autoimmune disease. The present sequence is a binding domain
CC   -immunoglobulin fusion protein-associated protein sequence. Note: The
CC   sequence data for this patent formed part of the printed specification
CC   and is also available in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not
CC   identified the sequences in the printed specification by their SEQ ID
CC   number therefore none of the sequences can be explicitly identified.
XX
SQ   Sequence 1153 AA;

  Query Match             99.9%;  Score 5868;  DB 7;  Length 1153;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              ||||||||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 7
AAR04136
ID   AAR04136 standard; protein; 1153 AA.
XX
AC   AAR04136;
XX
DT   09-SEP-2004  (revised)
DT   25-MAR-2003  (revised)
DT   07-SEP-1990  (first entry)
XX
DE   Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX
KW   Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
KW   non-specific defence system; integrin gene superfamily.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .16
FT                   /label= signal_peptide
FT   Modified-site   86. .88
FT                   /label= putative N-glycosylation site
FT   Modified-site   240. .242
FT                   /label= putative N-glycosylation site
FT   Modified-site   391. .393
FT                   /label= putative N-glycosylation site
FT   Modified-site   469. .471
FT                   /label= putative N-glycosylation site
FT   Modified-site   693. .695
FT                   /label= putative N-glycosylation site
FT   Modified-site   697. .699
FT                   /label= putative N-glycosylation site
FT   Modified-site   735. .737
FT                   /label= putative N-glycosylation site
FT   Modified-site   802. .804
FT                   /label= putative N-glycosylation site
FT   Modified-site   881. .883
FT                   /label= putative N-glycosylation site
FT   Modified-site   901. .903
FT                   /label= putative N-glycosylation site
FT   Modified-site   912. .914
FT                   /label= putative N-glycosylation site
FT   Modified-site   941. .943
FT                   /label= putative N-glycosylation site
FT   Modified-site   947. .949
FT                   /label= putative N-glycosylation site
FT   Modified-site   979. .981
FT                   /label= putative N-glycosylation site
```

```
FT   Modified-site    994. .996
FT                    /label= putative N-glycosylation site
FT   Modified-site    1022. .1024
FT                    /label= putative N-glycosylation site
FT   Modified-site    1045. .1047
FT                    /label= putative N-glycosylation site
FT   Modified-site    1051. .1053
FT                    /label= putative N-glycosylation site
FT   Modified-site    1076. .1078
FT                    /label= putative N-glycosylation site
FT   Region           1106. .1134
FT                    /label= putative_transmembrane_region
XX
PN   EP364690-A.
XX
PD   25-APR-1990.
XX
PF   17-AUG-1989;   89EP-00115159.
XX
PR   23-AUG-1988;   88US-00235353.
PR   09-MAR-1989;   89US-00321239.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Springer TA,  Corbi A;
XX
DR   WPI; 1990-125938/17.
DR   N-PSDB; AAQ04043.
XX
PT   New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating
PT   inflammation and viral infections, and in diagnosis.
XX
PS   Disclosure; Page ?; 3pp; English.
XX
CC   Mac-1 alpha subunit is involved in the response to inflammation, i.e.
CC   recognition of and migration to sites of inflammation. It also attaches
CC   to cellular substrates as part of this function making it useful in
CC   visualising endothelial tissue. Mac-1 is a member of the Integrin Gene
CC   superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25
CC   -MAR-2003 to correct PA field.)
CC
CC   Revised record issued on 09-SEP-2004 : Correction to feature table key
CC   and pages
XX
SQ   Sequence 1153 AA;

  Query Match             99.7%;  Score 5858;  DB 2;  Length 1153;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1132;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |||||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLLLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 8
ADM99589
ID   ADM99589 standard; protein; 1152 AA.
XX
AC   ADM99589;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Human integrin alphaM subunit precursor protein.
XX
KW   integrin alpha subunit; beta; antipsoriatic; thrombolytic; anticoagulant;
KW   osteopathic; cytostatic; immunosuppressive; antiinflammatory;
KW   neuroprotective; antisickling; immunotherapy; inflammatory;
KW   autoimmune disorder; thrombosis; cancer; osteoporosis;
KW   sickle cell anaemia; psoriasis; multiple sclerosis; human; precursor;
KW   alphaM.
XX
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   Misc-difference 965
FT                   /note= "Encoded by CCC"
XX
PN   WO2004007530-A2.
XX
PD   22-JAN-2004.
XX
PF   17-JUL-2003; 2003WO-US022301.
XX
PR   17-JUL-2002; 2002US-0396783P.
PR   17-JUL-2002; 2002US-0396790P.
PR   11-SEP-2002; 2002US-0410135P.
XX
PA   (BLOO-) CENT BLOOD RES INC.
XX
PI   Springer TA, Takagi J;
XX
DR   WPI; 2004-122877/12.
DR   N-PSDB; ADM99588.
XX
PT   Novel modified integrin protein having extracellular domains of integrin
PT   alpha and beta subunits or integrin alphaII and beta3 subunit, useful for
PT   treating integrin mediated disorders.
XX
PS   Disclosure; SEQ ID NO 4; 232pp; English.
XX
CC   The invention relates to a novel isolated or recombinant modified
CC   integrin protein having extracellular domains of integrin alpha and beta
CC   subunits where one of the subunits has one or more mutations, an altered
CC   surface feature or an amino acid substitution or internal deletion,
CC   extracellular domains of the integrin beta subunit that comprise a
CC   mutation that alters a non-cysteine residue to cysteine or extracellular
CC   domains of integrin alpha and beta subunits. The polypeptide of the
CC   invention demonstrates antipsoriatic, thrombolytic, anticoagulant,
CC   osteopathic, cytostatic, immunosuppressive, antiinflammatory,
CC   neuroprotective and antisickling activities and may be useful for
CC   immunotherapy in order to prevent or treat an integrin-mediated disorder
CC   such as an inflammatory disorder, an autoimmune disorder, thrombosis,
CC   cancer, osteoporosis, sickle cell anaemia, psoriasis and multiple
CC   sclerosis. The current sequence is that of the human integrin alphaM
CC   subunit precursor protein of the invention.
XX
SQ   Sequence 1152 AA;

  Query Match             99.6%;  Score 5852.5;  DB 8;  Length 1152;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1132;  Conservative    3;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              ||||||||||||||||||||||||||||||:|||:||| |||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 9
ADP12435
ID   ADP12435 standard; protein; 1152 AA.
XX
AC   ADP12435;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Protein encoded by mRNA of the invention #45.
XX
KW   transplant rejection; immune system; rheumatoid arthritis; lupus;
KW   inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX
OS   Homo sapiens.
XX
PN   WO2004042346-A2.
XX
PD   21-MAY-2004.
XX
PF   24-APR-2003; 2003WO-US012946.
```

XX
PR 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX
PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX
DR WPI; 2004-400724/37.
XX
PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
PS Claim 65; SEQ ID NO 2444; 1762pp; English.
XX
CC The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC protein that is encoded by the mRNA of the invention.
XX
SQ Sequence 1152 AA;

Query Match 99.5%; Score 5845.5; DB 8; Length 1152;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1131; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |||||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152
```

RESULT 10
ADP44061
ID ADP44061 standard; protein; 1163 AA.
XX
AC ADP44061;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human CD11C protein SEQ ID NO:14.
XX
KW HIV entry inhibitor; cell surface protein inhibitor; HIV infection;
KW anti-HIV; virucide; HIV uptake inhibitor; human; CD11C; integrin alpha X.
XX
OS Homo sapiens.
XX
PN WO2004053094-A2.
XX
PD 24-JUN-2004.
XX
PF 08-DEC-2003; 2003WO-US039208.
XX
PR 06-DEC-2002; 2002US-0431522P.
XX
PA (PPDP-) PPD DEV LP.
XX
PI Dunn SJ, Holzmayer TA;
XX
DR WPI; 2004-480928/45.
DR N-PSDB; ADP44060.
XX
PT Identifying an inhibitor of human immunodeficiency virus (HIV) entry into

```
PT   a human host cell useful for preventing and/or treating HIV infection, by
PT   identifying an inhibitor of a cell surface polypeptide such as CXCR-4.
XX
PS   Claim 1; SEQ ID NO 14; 133pp; English.
XX
CC   The present invention describes a method (M1) for identifying an
CC   inhibitor of HIV entry into a human host cell. (M1) comprises identifying
CC   an inhibitor of a cell surface polypeptide selected from CXCR-4 (352
CC   amino acids, SEQ ID NO:8, ADP44055), CCR4 (360 amino acids, SEQ ID NO:10,
CC   ADP44057), CCR7 (378 amino acids, SEQ ID NO:12, ADP44059), CD11C (1163
CC   amino acids, SEQ ID NO:14, ADP44061), CD47 (323 amino acids, SEQ ID
CC   NO:16, ADP44063), CD68 (354 amino acids, SEQ ID NO:18, ADP44065), CD69
CC   (199 amino acids, SEQ ID NO:20, ADP44067), CD74 (566 amino acids, SEQ ID
CC   NO:22, ADP44069), CSF3R (836 amino acids, SEQ ID NO:24, ADP44071), RARA
CC   (462 amino acids, SEQ ID NO:26, ADP44073), GABBR1 (578 amino acids, SEQ
CC   ID NO:28, ADP44075), P2X1 (150 amino acids, SEQ ID NO:30, ADP44077),
CC   HELO1 (299 amino acids, SEQ ID NO:32, ADP44079), GPRK6 (576 amino acids,
CC   SEQ ID NO:34, ADP44081), or PTK2B (1009 amino acids, SEQ ID NO:36,
CC   ADP44083). Also described: (1) an inhibitor of a cell surface polypeptide
CC   as described above in a human host cell preventing HIV entry into the
CC   human host cell; (2) a pharmaceutical composition comprising an inhibitor
CC   of (1) and a carrier; and (3) conferring resistance to HIV infection in
CC   an individual, by administering the pharmaceutical composition of (2). An
CC   inhibitor of HIV entry has anti-HIV and virucide activities, and can be
CC   used as an HIV uptake inhibitor. (M1) is useful for identifying
CC   protective compounds that inhibit entry of HIV into cells, useful for the
CC   prevention and/or treatment of HIV infection. The present sequence
CC   represents human integrin alpha X (CD11C), which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 1163 AA;

  Query Match             59.4%;  Score 3487;  DB 8;  Length 1163;
  Best Local Similarity   61.5%;  Pred. No. 4.6e-281;
  Matches  694;  Conservative  138;  Mismatches  291;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||    | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               |||| |||||||||||||:|||| |||||||||| |  | |: ||||| |    |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
               | : : ||::: || |||||||||   :|  |  ||  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              : ||||:||: : || ||:  : || | |:||| :: |:  | : : |||::| ||||||
Db        198 QTHPTFEEFRRSSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAAKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
               |||:| || | |:|||| ||  |:||| |||| ||::  | :||| ||||| ::|:|:|
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:|||||||||:| |||||| ||:||||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:|||||:  : |   ||||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |    | | |||||:||||||||||||:|||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDSDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || |  |||||||||||||:|| |||||||:|||||||| ||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :|||  :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:|| |||| |||  |||||| |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAADAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |     |||| ||| ||   :::|| |||| || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| ||:||   ||||||: | |  | |  |:|||:||||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| |  | :::  |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :      : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :|  ::| || |||||||:|||||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143


RESULT 11
ADN02004
ID   ADN02004 standard; protein; 1163 AA.
XX
AC   ADN02004;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Human inflammatory bowel disease marker - CD11c protein.
XX
KW   marker gene; inflammatory bowel disease; IBD; FcgammaR IIIa;
KW   FcgammaR IIIb; Mig; NRG-2; hexokinase; HM74; REG III; LPAP; Mip-1(beta);
KW   L-selectin; EGFL6; IDO; IL-8; CD11c; TLR2; ulcerative colitis;
KW   Crohn's disease.
XX
OS   Homo sapiens.
XX
PN   JP2004065120-A.
XX
PD   04-MAR-2004.
XX
PF   07-AUG-2002; 2002JP-00229705.
XX
PR   07-AUG-2002; 2002JP-00229705.
XX
PA   (SUMU ) SUMITOMO SEIYAKU KK.
XX
DR   WPI; 2004-209124/20.
DR   N-PSDB; ADN01990.
XX
PT   Novel marker gene of inflammatory bowel disease (IBD) comprising base
PT   sequence of e.g., Fc gamma R IIIa, Fc gamma R IIIb, Mig, NRG-2,
PT   hexokinase 3, HM74, CD11c, TLR2 gene, useful for screening therapeutic
```

PT agent for IBD.
XX
PS Example; SEQ ID NO 27; 151pp; Japanese.
XX
CC The invention comprises marker genes for inflammatory bowel disease (IBD)
CC - FcgammaR IIIa, FcgammaR IIIb, Mig, NRG-2, hexokinase, HM74, REG III,
CC LPAP, Mip-1(beta), L-selectin, EGFL6, IDO, IL-8, CD11c, and TLR2 genes.
CC The DNA and protein sequences of the invention are useful in the
CC detection and treatment of IBD (e.g. ulcerative colitis and Crohn's
CC disease). The present amino acid sequence represents an IBD marker
CC protein of the invention.
XX
SQ Sequence 1163 AA;

Query Match 59.1%; Score 3473; DB 8; Length 1163;
Best Local Similarity 61.3%; Pred. No. 6.9e-280;
Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||     |: :: ||| ||||    | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               |||| |||||||||||||:|||| ||||||||||  |  | |: |||||  |     | |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
               | : : ||::: || |||||||||   :|  |  ||  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              : ||||:||:   || ||:  : || | |:||| :: |:  | : :  |||::| ||||||
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              |||:| || | |:|||| ||  |:||| |||| ||::  | :||| ||||| ::|:|:|
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:|||||||||:| |||||| ||:||||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |   |||||||||||||  |    :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | |||||:|||||||||||||::|||||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||||| ||||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||  |  |||||||||||||:|| |||||||:|||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :|||   :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
               ||||| ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || |||||:||  ||| |||  ||||||| |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | |  | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854
```

```
Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |    |||| ||| ||   :::||  |||| || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| ||:||  |||||: | |  | |  |:|||:|||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | |||  :|   :|:  :| |  | ::: |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :     : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 12
ADQ17510
ID ADQ17510 standard; protein; 1163 AA.
XX
AC ADQ17510;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 327.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 327; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 1163 AA;

Query Match 59.1%; Score 3473; DB 8; Length 1163;
Best Local Similarity 61.3%; Pred. No. 6.9e-280;
Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

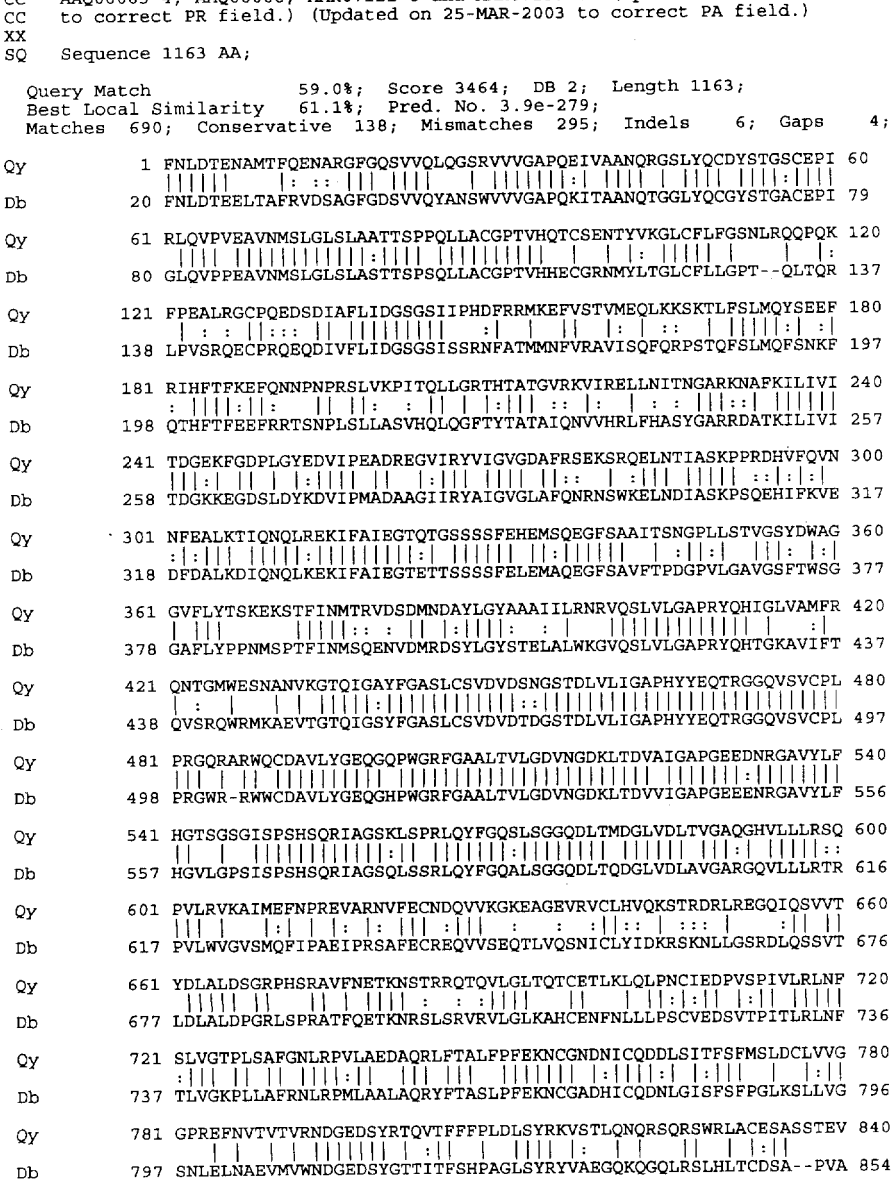
RESULT 13
AAR07120
ID AAR07120 standard; protein; 1163 AA.
XX
AC AAR07120;
XX
DT 25-MAR-2003 (revised)
DT 05-FEB-1991 (first entry)
XX
DE p150.95 alpha subunit encoded by clone lambdaX47.
XX
KW p150,95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;
KW rhinovirus.
XX
OS Synthetic.
XX

FH	Key	Location/Qualifiers
FT	Region	1. .19
FT		/label= signal peptide
FT	Region	20. .44
FT		/label= N-terminus
FT	Modified-site	61. .63
FT		/label= glycosylation site
FT	Modified-site	89. .91
FT		/label= glycosylation site
FT	Modified-site	385. .387
FT		/label= glycosylation site
FT	Modified-site	392. .394
FT		/label= glycosylation site
FT	Modified-site	697. .699
FT		/label= glycosylation site
FT	Modified-site	735. .737
FT		/label= glycosylation site
FT	Modified-site	899. .901
FT		/label= glycosylation site
FT	Modified-site	904. .906
FT		/label= glycosylation site
FT	Modified-site	939. .941
FT		/label= glycosylation site
FT	Modified-site	1050. .1052
FT		/label= glycosylation site
FT	Domain	1108. .1133
FT		/label= transmembrane

XX
PN WO9010646-A.
XX
PD 20-SEP-1990.
XX
PF 09-MAR-1990; 90WO-US001257.
XX
PR 09-MAR-1990; 90WO-US001257.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Corbi AA, Springer TA;
XX
DR WPI; 1990-304985/40.
DR N-PSDB; AAQ06068.
XX
PT Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-
PT unit of p150.95 cell surface adhesion receptor, opt. together with a beta
PT chain of CD-18 family.
XX
PS Disclosure; Fig 3; 59pp; English.
XX

CC Clone lambda X47 was isolated from a cDNA library constructed from total
CC RNA extracted from phorbol myristate acetate stimulated HL-60
CC myelomonocytic cells. The library was screened with oligonucleotide
CC probes based on tryptic peptide fragments of p150.95. The sequence can be
CC attached to appropriate control elements and expressed in prokaryotic and
CC eukaryotic cells. The protein can be used to treat or prevent rhinoviral
CC infection because it interacts with ICAM-1 and inhibits cell-virus
CC attachment. It can also be used as an anti-inflammatory agent. See also
CC AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003
CC to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1163 AA;

Query Match 59.0%; Score 3464; DB 2; Length 1163;
Best Local Similarity 61.1%; Pred. No. 3.9e-279;
Matches 690; Conservative 138; Mismatches 295; Indels 6; Gaps 4;



RESULT 14
AAW65091
ID AAW65091 standard; protein; 1163 AA.
XX
AC AAW65091;
XX
DT 28-SEP-1998 (first entry)
XX
DE Human Beta-integrin CD11c subunit protein.
XX
KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN US5728533-A.
XX
PD 17-MAR-1998.
XX
PF 07-JUN-1995; 95US-00485618.
XX
PR 23-DEC-1993; 93US-00173497.
PR 05-AUG-1994; 94US-00286889.
PR 21-DEC-1994; 94US-00362652.
XX
PA (ICOS-) ICOS CORP.
XX
PI Van Der Vieren M, Gallatin WM;
XX
DR WPI; 1998-206565/18.
XX
PT Screening assay for modulators of integrin binding - using immobilised or
PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.
XX
PS Example 5; Fig 1A-D; 106pp; English.
XX
CC This sequence represents a human beta-integrin CD11c subunit which is
CC used to describe a method for identifying compounds that modulate the
CC interaction of the beta-integrin alpha-d subunit with a binding partner
CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha
CC -d binding partner, one of which is immobilised and the other of which is
CC labelled, in the presence of a test compound, and determining if the
CC compound affects binding between the alpha-d polypeptide and alpha-d
CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment
CC comprising the cytoplasmic, transmembrane or extracellular domain of
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat
CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC and rheumatoid arthritis
XX

SQ Sequence 1163 AA;

Query Match 58.7%; Score 3450; DB 2; Length 1163;
Best Local Similarity 61.0%; Pred. No. 5.7e-278;
Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAIKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        318 DFDALKDIQNQLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        438 QVSRQWRMKAEVIGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALAPGRLSPRAIFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTC--CSAPVG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYIVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEIIFDTSVYSQLPGQEAFMRAQTI 1094

Qy       1080 TKVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPIPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 15
AAB07361
ID AAB07361 standard; protein; 1163 AA.
XX
AC AAB07361;
XX
DT 17-JAN-2001 (first entry)
XX
DE Human CD11c protein sequence.
XX
KW Human; macrophage infiltration inhibition; alpha_d integrin;
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW rheumatoid arthritis; central nervous system injury; CD11c.
XX
OS Homo sapiens.
XX
PN WO200029446-A1.
XX
PD 25-MAY-2000.
XX
PF 16-NOV-1999; 99WO-US027139.
XX
PR 16-NOV-1998; 98US-00193043.
PR 08-JUL-1999; 99US-00350259.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gallatin MW, Van Der Vieren M;
XX
DR WPI; 2000-387751/33.
XX
PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous system
PT injury sites.
XX
PS Example 5; Fig 1; 270pp; English.
XX
CC Integrins are a class of membrane-associated molecules that participate
CC in cellular adhesion. Integrins are made up of an alpha subunit and a
CC beta subunit. One class of human integrins are restricted to expression
CC in white blood cells and have a common beta2 subunit: the leukocyte
CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC have an important role in immune and inflammatory responses. The present
CC protein sequence is the human integrin alpha subunit CD11c. This sequence
CC was used in an alignment to identify a novel beta2 integrin alpha
CC subunit: alpha_d (AAA60014 and AAB07359). The present sequence has
CC approximately 66% identity to the protein sequence of alpha_d. The
CC Alpha_d gene and protein may be useful in therapy for diseases linked to
CC aberrant alpha_d function e.g. Type I diabetes, atherosclerosis, multiple
CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
CC (LAD). In addition, anti-alpha_d monoclonal antibodies may be used in the
CC inhibition of macrophage infiltration at the site of a central nervous
CC system injury. The monoclonal antibodies can also be used to detect and
CC diagnose Crohn's disease
XX
SQ Sequence 1163 AA;

Query Match 58.7%; Score 3450; DB 3; Length 1163;
Best Local Similarity 61.0%; Pred. No. 5.7e-278;

Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAIKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        318 DFDALKDIQNQLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        438 QVSRQWRMKAEVIGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALAPGRLSPRAIFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALPFPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTC--CSAPVG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYIVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEIIFDTSVYSQLPGQEAFMRAQTI 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPIPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

Search completed: November 9, 2004, 12:08:29
Job time : 108.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 12:17:49 ; Search time 97.5 Seconds
(without alignments)
4119.157 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5875	100.0	1137	10	US-09-902-481A-6	Sequence 6, Appli
2	5868	99.9	1153	9	US-09-350-259-3	Sequence 3, Appli
3	5868	99.9	1153	10	US-09-902-481A-1	Sequence 1, Appli
4	5868	99.9	1153	10	US-09-891-943-3	Sequence 3, Appli
5	5868	99.9	1153	14	US-10-144-259-30	Sequence 30, Appl
6	5868	99.9	1153	14	US-10-207-655-176	Sequence 176, App
7	5855	99.7	1137	10	US-09-902-481A-5	Sequence 5, Appli
8	5852.5	99.6	1152	9	US-09-945-265-4	Sequence 4, Appli
9	5845	99.5	1137	10	US-09-902-481A-4	Sequence 4, Appli
10	5832	99.3	1137	10	US-09-902-481A-3	Sequence 3, Appli
11	3473	59.1	1163	14	US-10-116-275-204	Sequence 204, App
12	3450	58.7	1163	9	US-09-350-259-4	Sequence 4, Appli
13	3450	58.7	1163	10	US-09-891-943-4	Sequence 4, Appli
14	3417	58.2	1161	9	US-09-350-259-2	Sequence 2, Appli
15	3417	58.2	1161	10	US-09-891-943-2	Sequence 2, Appli
16	3401.5	57.9	1161	9	US-09-350-259-99	Sequence 99, Appl
17	3401.5	57.9	1161	10	US-09-891-943-99	Sequence 99, Appl
18	3239.5	55.1	1161	9	US-09-350-259-55	Sequence 55, Appl
19	3239.5	55.1	1161	10	US-09-891-943-55	Sequence 55, Appl
20	3228	54.9	1161	9	US-09-350-259-53	Sequence 53, Appl
21	3228	54.9	1161	10	US-09-891-943-53	Sequence 53, Appl
22	3218.5	54.8	1151	9	US-09-350-259-37	Sequence 37, Appl
23	3218.5	54.8	1151	10	US-09-891-943-37	Sequence 37, Appl
24	3207	54.6	1155	9	US-09-350-259-46	Sequence 46, Appl
25	3207	54.6	1155	10	US-09-891-943-46	Sequence 46, Appl
26	1848	31.5	369	13	US-10-087-192-1212	Sequence 1212, Ap
27	1552.5	26.4	1170	9	US-09-945-265-2	Sequence 2, Appli
28	1548.5	26.4	1170	15	US-10-261-164-1	Sequence 1, Appli
29	1525	26.0	1223	16	US-10-408-765A-295	Sequence 295, App
30	1350.5	23.0	1086	16	US-10-408-765A-1871	Sequence 1871, Ap
31	1229.5	20.9	494	9	US-09-350-259-103	Sequence 103, App
32	1229.5	20.9	494	10	US-09-891-943-103	Sequence 103, App
33	1157.5	19.7	413	9	US-09-350-259-101	Sequence 101, App
34	1157.5	19.7	413	10	US-09-891-943-101	Sequence 101, App
35	1149	19.6	1179	14	US-10-177-550-2	Sequence 2, Appli
36	1149	19.6	1179	14	US-10-173-551-2	Sequence 2, Appli
37	1102.5	18.8	1151	10	US-09-984-130-103	Sequence 103, App
38	1102.5	18.8	1151	10	US-09-836-353A-103	Sequence 103, App
39	1102.5	18.8	1179	10	US-09-918-715-250	Sequence 250, App
40	1102.5	18.8	1179	17	US-10-474-794-250	Sequence 250, App
41	1101	18.7	1188	14	US-10-291-265-810	Sequence 810, App
42	1097	18.7	1188	14	US-10-291-265-338	Sequence 338, App
43	1093.5	18.6	1189	10	US-09-984-130-35	Sequence 35, Appl
44	1093.5	18.6	1189	10	US-09-836-353A-35	Sequence 35, Appl
45	1093.5	18.6	1189	15	US-10-262-839-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-902-481A-6
; Sequence 6, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6

Query Match 100.0%; Score 5875; DB 10; Length 1137;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
```

```
Qy       121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db       121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180

Qy       181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db       181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240

Qy       241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db       241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300

Qy       301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db       301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy       361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db       361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy       421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db       421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy       481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db       481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy       541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db       541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy       601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db       601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy       661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db       661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy       721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db       721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy       781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db       781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy       841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db       841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy       901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db       901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy       961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db       961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy      1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db      1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy      1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db      1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

RESULT 2
US-09-350-259-3
; Sequence 3, Application US/09350259

```
; Patent No. US20020062008A1
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT: Van der Vieren, Monica
;  TITLE OF INVENTION: No. US20020062008A1el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/350,259
;  CURRENT FILING DATE: 1999-07-08
;  EARLIER APPLICATION NUMBER: 09/193,043
;  EARLIER FILING DATE: 1998-11-16
;  EARLIER APPLICATION NUMBER: 08/173,497
;  EARLIER FILING DATE: 1993-12-23
;  EARLIER APPLICATION NUMBER: 08/286,889
;  EARLIER FILING DATE: 1994-08-05
;  EARLIER APPLICATION NUMBER: 08/362,652
;  EARLIER FILING DATE: 1994-12-21
;  EARLIER APPLICATION NUMBER: 08/943,363
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-350-259-3

  Query Match           99.9%;  Score 5868;  DB 9;  Length 1153;
  Best Local Similarity 99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db        17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy        61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db        77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy       121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db       137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy       181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db       197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy       241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db       257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy       301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db       317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy       361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db       377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy       421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db       437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy       481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db       497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy       541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db       557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy       601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db       617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676
```

```
Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 3
US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:
US-09-902-481A-1

Query Match 99.9%; Score 5868; DB 10; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 4
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1

```
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT:  Van der Vieren, Monica
;  TITLE OF INVENTION: No. US20030077278A1el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/891,943
;  CURRENT FILING DATE:  2001-06-26
;  PRIOR APPLICATION NUMBER: 09/193,043
;  PRIOR FILING DATE: 1998-11-16
;  PRIOR APPLICATION NUMBER: 08/286,889
;  PRIOR FILING DATE: 1994-08-05
;  PRIOR APPLICATION NUMBER: 08/362,652
;  PRIOR FILING DATE: 1994-12-21
;  PRIOR APPLICATION NUMBER: 08/943,363
;  PRIOR FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-891-943-3

  Query Match             99.9%;  Score 5868;  DB 10;  Length 1153;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |||||||||||||||||||||||||||||||||:|||:||| |||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 5
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
;  APPLICANT: Arnaout, M. Amin
;  APPLICANT:  Li, Rui
;  APPLICANT:  Xiong, Jian-Ping
;  TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: 00786-548001
;  CURRENT APPLICATION NUMBER: US/10/144,259
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: US 09/758,493
;  PRIOR FILING DATE: 2001-01-11
;  PRIOR APPLICATION NUMBER: US 60/221,950
;  PRIOR FILING DATE: 2000-07-31
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-144-259-30

  Query Match             99.9%;  Score 5868;  DB 14;  Length 1153;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |||||||||||||||||||||||||||||||||:|||:||| |||||||||||||:||
```

```
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256
Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316
Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376
Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436
Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496
Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616
Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676
Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736
Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796
Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856
Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916
Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976
Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036
Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096
Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 6
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 99.9%; Score 5868; DB 14; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76
Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136
Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196
Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256
Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316
Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376
Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436
Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496
Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616
Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676
Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736
Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796
Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856
Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916
```

```
Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 7
US-09-902-481A-5
; Sequence 5, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-5

Query Match 99.7%; Score 5855; DB 10; Length 1137;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

RESULT 8
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; TITLE OF INVENTION: DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: CBN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4

Query Match 99.6%; Score 5852.5; DB 9; Length 1152;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1132; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              ||||||||||||||||||||||||||||||||:|||:||| |||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152
```

RESULT 9
US-09-902-481A-4
; Sequence 4, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4

Query Match 99.5%; Score 5845; DB 10; Length 1137;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |||||||||||||||||||:|||||||||||||:||| |||||||||||||  ::
Db        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              |||||||||||||||||| ||||||||||||||||||||||||||||:|||||||||:|
Db        241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
```

```
Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

RESULT 10
US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 99.3%; Score 5832; DB 10; Length 1137;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1122; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFSLMQYSEEF 180

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db        181 RIHFTFKEFQNNPNPRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILFLL 240

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db        241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        301 NFEALKTVQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

RESULT 11
US-10-116-275-204
; Sequence 204, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 204
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-204

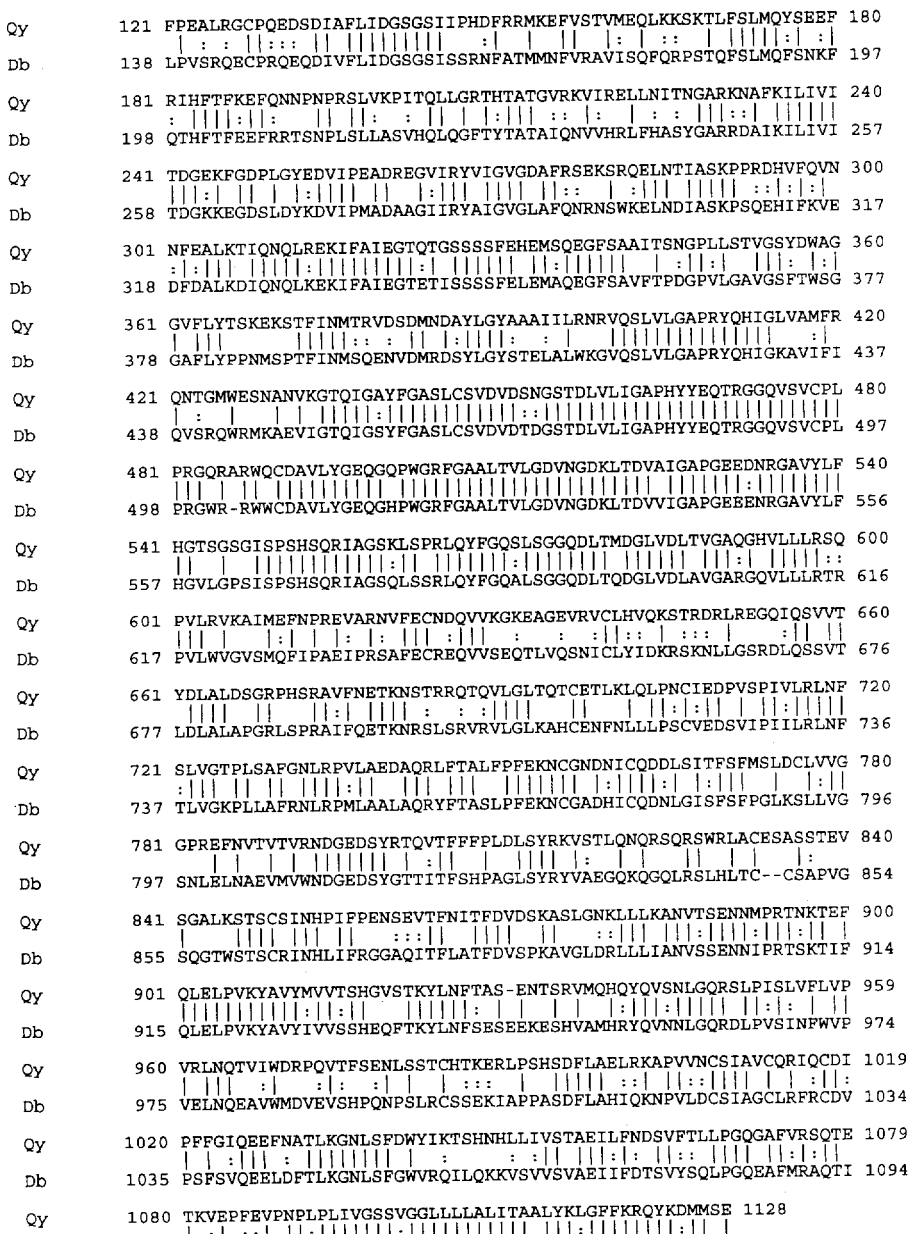
Query Match 59.1%; Score 3473; DB 14; Length 1163;
Best Local Similarity 61.3%; Pred. No. 6e-284;
Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||| |: :: ||| |||| | |||||||:| |||| | |||| ||||:|||| |
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||:|||| |||||||||| | | |: ||||| | | |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || |||||||||   :| | || |: | :: | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              : ||||:||:   || ||:  : || | |:||| :: |: | : : |||::| ||||||
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              |||:| || | |:|||| || |:||| |||| ||::  | :||| ||||| ::|:|:|
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              :|:||| |||||:|||||||||:| |||||| ||:|||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||       |||||:: : || |:||||:  : |  ||||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |   | | ||||||:||||||||||||::||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| | || |||||||||| |||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || |  |||||||||||||:|| |||||||:|||||||| |||||| |||:| |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||| |   |:| | |: |: ||| :|||  :   :  :||:: | ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||| ||   || | |||| :  : :|||| || | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              :||| || || ||||:||  ||| ||| |||||| |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | | | | ||||||| | :||  |  |||| |:  | |  || | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              |  |||| ||| ||  :::|| |||| || ||::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||| ||:||  ||||||: | |  | |  |:|||:||||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | ||| :|  :|:  :| | | ::: |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :     : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
              | :| ::| || |||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 12
US-09-350-259-4
; Sequence 4, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-4

Query Match 58.7%; Score 3450; DB 9; Length 1163;
Best Local Similarity 61.0%; Pred. No. 5.3e-282;
Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||||| |: :: ||| |||| | |||||||:|:|||| | |||| ||||:|||| |
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| |||||||||||:|||| |||||||||| | | |: ||||| | | |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137
```



RESULT 13
US-09-891-943-4
; Sequence 4, Application US/09891943

; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-4

Query Match 58.7%; Score 3450; DB 10; Length 1163;
Best Local Similarity 61.0%; Pred. No. 5.3e-282;
Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;

```
Db        677 LDLALAPGRLSPRAIFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTC--CSAPVG 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959

Db        915 QLELPVKYAVYIVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019

Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079

Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEIIFDTSVYSQLPGQEAFMRAQTI 1094

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128

Db       1095 TVLEKYKVHNPIPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

RESULT 14
US-09-350-259-2
; Sequence 2, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-2

Query Match 58.2%; Score 3417; DB 9; Length 1161;
Best Local Similarity 59.8%; Pred. No. 3.3e-279;
Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQTGRLYDCAAATGMCQPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-EIIQT 135

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180

Db        136 VPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFEGTDTLFALMQYSNLL 195

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240

Db        196 KIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHHKNGARKSAKKILIVI 255

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300

Db        256 TDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELNTISSAPPQDHVFKVD 315

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        316 NFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGSFSWSG 375

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        376 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVLGAPRYQHTGKAVIFT 435

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        436 QVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPHYYEQTRGGQVSVCPL 495

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        496 PRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAIGAPGEQENRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLLRSL 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        616 PVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        674 FDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCVEDVVSPIILHLNF 733

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        734 SLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDLGVTLSFSGLQTLTVG 793

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        794 SSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        853 EG-LRSSRCSVNHPIFHEGSNGTFIVTFDVSYKATLGDRMLMRASASSENNKASSSKATF 911

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNF-TASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959

Db        912 QLELPVKYAVYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQRDLAISINFWVP 971

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019

Db        972 VLLNGVAVWDVVMEAPSQSL--PCVSERKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1029

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079

Db       1030 PSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQLPGQEAFMRAQME 1089

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128

Db       1090 MVLEEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRHYKEMLED 1138
```

RESULT 15
US-09-891-943-2
; Sequence 2, Application US/09891943
; Publication No. US20030077278A1

; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-2

Query Match 58.2%; Score 3417; DB 10; Length 1161;
Best Local Similarity 59.8%; Pred. No. 3.3e-279;
Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              |||| |    |||:| ||||||||  |||:||||| |:||||| | || |  :|| |:||
Db         17 FNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQTGRLYDCAAATGMCQPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
               | :  |||||||||:|||:|:  :|||||||:|: | ||:| || | | ||   :  |
Db         77 PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-EIIQT 135

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              |:|   || :: || |||||||||  :|| :|| ||  || | : : |||:|||||
Db        136 VPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFEGTDTLFALMQYSNLL 195

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              :||||| :|: :|: :||| || || | | ||||:  |: :| :  |||||:| ||||||
Db        196 KIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHHKNGARKSAKKILIVI 255

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              |||:|: ||| | ||||:|:: |:||| |||| ||:   :|||||||:| ||:||||:|:
Db        256 TDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELNTISSAPPQDHVFKVD 315

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              || || :|| ||:|||:|:|||: :||||:|||||||||| |:| :|  |  |||: |:|
Db        316 NFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGSFSWSG 375

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              | |||      |||||:: : || |:||||:  : |   ||:||||||||| |    :|
Db        376 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVLGAPRYQHTGKAVIFT 435

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              | :  |    | | |||||:|||||||||||||:|||||:||||||||||||||||||||
Db        436 QVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPHYYEQTRGGQVSVCPL 495

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||| :|||||| |||| ||||||||||||||| ||| |||||||||::|||||||
Db        496 PRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAIGAPGEQENRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              || | |||||||||||| |:||||||||||:|||||||| |||:|| |||:| ||||||
Db        556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLLRSL 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              |||:|    | |:| |||: |: | ::     |||:  ||| :|||: |:|  | ||| |
Db        616 PVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              :||||| ||  |||:|||||| |  : : |||   |||||| ||:|:|| ||||:| |||
Db        674 FDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCVEDVVSPIILHLNF 733

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              |||  |: :  |||||||  :| ||||  |||||||  | :|: || :| ||  |  | ||
Db        734 SLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDLGVTLSFSGLQTLTVG 793

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              | || ||| | ||||| | |: ::|   ||:|:||  | |  | : |||||:   ||
Db        794 SSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              | |:|: ||:|||| | |  || :|||  ||:||:::|::|: :||||   ::|  |
Db        853 EG-LRSSRCSVNHPIFHEGSNGTFIVTFDVSYKATLGDRMLMRASASSENNKASSSKATF 911

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNF-TASENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
              |||||||||||| :::    |||| || |: |   :  :|:|:|:|| || | ||: | ||
Db        912 QLELPVKYAVYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQRDLAISINFWVP 971

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
              | ||   :||      |::|   | :: : | ||||| :: ::|:::|||| | : :||:
Db        972 VLLNGVAVWDVVMEAPSQSL--PCVSERKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1029

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
              | | :||| : |||||||| |  :|    :|:|| ||| |: ||:: |||| ||:|:| |
Db       1030 PSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQLPGQEAFMRAQME 1089

Qy       1080 TKVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
                :|  || | :|:|:||||| |||||||| ||||||| ||:|: :
Db       1090 MVLEEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRHYKEMLED 1138
```

Search completed: November 9, 2004, 12:47:03
Job time : 102.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:57 ; Search time 23 Seconds
(without alignments)
3278.416 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5868	99.9	1153	1	US-08-173-497-3	Sequence 3, Appli
2	5868	99.9	1153	1	US-08-286-889-3	Sequence 3, Appli
3	5868	99.9	1153	1	US-08-485-618-3	Sequence 3, Appli
4	5868	99.9	1153	1	US-08-362-652-3	Sequence 3, Appli
5	5868	99.9	1153	2	US-08-605-672-3	Sequence 3, Appli
6	5868	99.9	1153	2	US-08-482-293A-3	Sequence 3, Appli
7	5868	99.9	1153	2	US-08-943-363-3	Sequence 3, Appli
8	5868	99.9	1153	3	US-09-193-043-3	Sequence 3, Appli
9	5868	99.9	1153	4	US-09-688-307A-3	Sequence 3, Appli
10	5868	99.9	1153	4	US-09-350-259-3	Sequence 3, Appli
11	5837.5	99.4	1152	2	US-08-476-062A-43	Sequence 43, Appl
12	5837.5	99.4	1152	5	PCT-US96-01314-43	Sequence 43, Appl
13	5837.5	99.4	1152	6	5424399-2	Patent No. 5424399
14	3473	59.1	1163	2	US-08-476-062A-44	Sequence 44, Appl
15	3473	59.1	1163	5	PCT-US96-01314-44	Sequence 44, Appl
16	3450	58.7	1163	1	US-08-173-497-4	Sequence 4, Appli
17	3450	58.7	1163	1	US-08-286-889-4	Sequence 4, Appli
18	3450	58.7	1163	1	US-08-485-618-4	Sequence 4, Appli
19	3450	58.7	1163	1	US-08-362-652-4	Sequence 4, Appli
20	3450	58.7	1163	2	US-08-605-672-4	Sequence 4, Appli
21	3450	58.7	1163	2	US-08-482-293A-4	Sequence 4, Appli
22	3450	58.7	1163	2	US-08-943-363-4	Sequence 4, Appli
23	3450	58.7	1163	3	US-09-193-043-4	Sequence 4, Appli
24	3450	58.7	1163	4	US-09-688-307A-4	Sequence 4, Appli
25	3450	58.7	1163	4	US-09-350-259-4	Sequence 4, Appli
26	3417	58.2	1161	1	US-08-173-497-2	Sequence 2, Appli
27	3417	58.2	1161	1	US-08-286-889-2	Sequence 2, Appli
28	3417	58.2	1161	1	US-08-485-618-2	Sequence 2, Appli
29	3417	58.2	1161	1	US-08-362-652-2	Sequence 2, Appli
30	3417	58.2	1161	2	US-08-605-672-2	Sequence 2, Appli
31	3417	58.2	1161	2	US-08-482-293A-2	Sequence 2, Appli
32	3417	58.2	1161	2	US-08-943-363-2	Sequence 2, Appli
33	3417	58.2	1161	3	US-09-193-043-2	Sequence 2, Appli
34	3417	58.2	1161	4	US-09-688-307A-2	Sequence 2, Appli
35	3417	58.2	1161	4	US-09-350-259-2	Sequence 2, Appli
36	3401.5	57.9	1161	1	US-08-485-618-99	Sequence 99, Appl
37	3401.5	57.9	1161	2	US-08-605-672-99	Sequence 99, Appl
38	3401.5	57.9	1161	2	US-08-482-293A-99	Sequence 99, Appl
39	3401.5	57.9	1161	2	US-08-943-363-99	Sequence 99, Appl
40	3401.5	57.9	1161	3	US-09-193-043-99	Sequence 99, Appl
41	3401.5	57.9	1161	4	US-09-688-307A-99	Sequence 99, Appl
42	3401.5	57.9	1161	4	US-09-350-259-99	Sequence 99, Appl
43	3239.5	55.1	1161	3	US-09-193-043-55	Sequence 55, Appl
44	3239.5	55.1	1161	4	US-09-688-307A-55	Sequence 55, Appl
45	3239.5	55.1	1161	4	US-09-350-259-55	Sequence 55, Appl

ALIGNMENTS

```
RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van Der Vieren, Monica
;    TITLE OF INVENTION:  No. 5437958el Human  2 Integrin Alpha
;    TITLE OF INVENTION:  Subunit
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 S. Wacker Drive, 6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/173,497
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5437958and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/31363
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-173-497-3

  Query Match           99.9%;  Score 5868;  DB 1;  Length 1153;
  Best Local Similarity  99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76
```

```
Qy        61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db        77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy       121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db       137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy       181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db       197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy       241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db       257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy       301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db       317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy       361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db       377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy       421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db       437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy       481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db       497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy       541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db       557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy       601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db       617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy       661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db       677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy       721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db       737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy       781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db       797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy       841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db       857 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy       901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db       917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy       961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db       977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy      1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db      1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy      1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db      1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Mich
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5470953el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,889
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  P38,659
;      REFERENCE/DOCKET NUMBER:  27866/32168
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-286-889-3

  Query Match           99.9%;  Score 5868;  DB 1;  Length 1153;
  Best Local Similarity  99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376
```

```
Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 3
US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5728533el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,618
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32797
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-485-618-3

  Query Match             99.9%;  Score 5868;  DB 1;  Length 1153;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              ||||||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 4
US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5766850
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5766850el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/362,652
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32391
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1153 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-362-652-3

  Query Match             99.9%;  Score 5868;  DB 1;  Length 1153;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |||||||||||||||||||||||||||||||:|||:||| ||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 5
US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-672-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |||||||||||||||||||||||||||||||||||:|||:||| |||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036
```

```
Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 6
US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-293A-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              |||||||||||||||||||||||||||||||||:|||:||| ||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,363
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-943-363-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy    1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
Db   17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy   61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
Db   77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy  121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
Db  137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy  181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
Db  197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy  241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
Db  257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy  301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db  317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy  361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
Db  377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy  421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
Db  437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy  481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db  497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy  541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db  557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy  601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db  617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy  661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db  677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy  721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db  737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy  781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db  797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy  841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db  857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy  901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
Db  917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy  961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
Db  977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy 1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
Db 1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy 1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
Db 1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03

```
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-193-043-3

  Query Match            99.9%;  Score 5868;  DB 3;  Length 1153;
  Best Local Similarity  99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              ||||||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 SGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153


RESULT 9
US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT:  Van der Vieren, Monica
;  TITLE OF INVENTION: No. 6432404el Human Beta-2
;  FILE REFERENCE: 27866/36646
;  CURRENT APPLICATION NUMBER: US/09/688,307A
;  CURRENT FILING DATE:  2000-10-13
;  PRIOR APPLICATION NUMBER: 09/193,043
;  PRIOR FILING DATE: 1998-11-16
;  PRIOR APPLICATION NUMBER: 08/605,672
;  PRIOR FILING DATE: 1996-02-22
;  PRIOR APPLICATION NUMBER: 08/173,497
;  PRIOR FILING DATE: 1993-12-23
;  PRIOR APPLICATION NUMBER: 08/286,889
;  PRIOR FILING DATE: 1994-08-05
;  PRIOR APPLICATION NUMBER: 08/362,652
;  PRIOR FILING DATE: 1994-12-21
;  PRIOR APPLICATION NUMBER: 08/943,363
;  PRIOR FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-688-307A-3

  Query Match            99.9%;  Score 5868;  DB 4;  Length 1153;
  Best Local Similarity  99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              ||||||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 10
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
;  APPLICANT: Gallatin, Michael W.
;  APPLICANT: Van der Vieren, Monica
;  TITLE OF INVENTION: No. 6620915el Human 2
;  FILE REFERENCE: 27866/35004
;  CURRENT APPLICATION NUMBER: US/09/350,259
;  CURRENT FILING DATE: 1999-07-08
;  EARLIER APPLICATION NUMBER: 09/193,043
;  EARLIER FILING DATE: 1998-11-16
;  EARLIER APPLICATION NUMBER: 08/173,497
;  EARLIER FILING DATE: 1993-12-23
;  EARLIER APPLICATION NUMBER: 08/286,889
;  EARLIER FILING DATE: 1994-08-05
;  EARLIER APPLICATION NUMBER: 08/362,652
;  EARLIER FILING DATE: 1994-12-21
;  EARLIER APPLICATION NUMBER: 08/943,363
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-350-259-3

  Query Match             99.9%;  Score 5868;  DB 4;  Length 1153;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180

Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240

Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300

Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        617 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        677 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        737 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        797 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 856

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
```

```
Db        857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 916

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        917 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 976

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1036

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1096

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137

Db       1097 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1153
```

```
RESULT 11
US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
;  GENERAL INFORMATION:
;    APPLICANT:  Arnaout, M. Amin
;    TITLE OF INVENTION:  CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
;    TITLE OF INVENTION:  RESPONSES WITH BETA2 INTEGRINS
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,062A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/216,081
;      FILING DATE:  21-MAR-1994
;      APPLICATION NUMBER:  07/637,830
;      FILING DATE:  04-JAN-1991
;      APPLICATION NUMBER:  07/539,842
;      FILING DATE:  18-JUN-1990
;      APPLICATION NUMBER:  07/212,573
;      FILING DATE:  28-JUN-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Freeman, John W.
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  00786/068003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1152 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-476-062A-43

  Query Match            99.4%;  Score 5837.5;  DB 2;  Length 1152;
  Best Local Similarity  99.4%;  Pred. No. 0;
  Matches 1130;  Conservative    3;  Mismatches    3;  Indels    1;  Gaps    1;
```

```
Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180

Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240

Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300

Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360

Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420

Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480

Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600

Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660

Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720

Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780

Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840

Db        796 GPRESNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900

Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960

Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020

Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080

Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
```

```
                                                                                                                    |||||||
```

```
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1130;  Conservative    3;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 76

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 136

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
              ||||||||||||||||||||||||||||||||:|||:||| ||||||||||||||:||
Db        197 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 436

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 496

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 PRG-RARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 615

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 675

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 735

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GPRESNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 855

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 915

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSLPISLVFLVPV 960
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        916 QLELPVKYAVYMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLGQRSPPISLVFLVPV 975

Qy        961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDIP 1035

Qy       1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 FFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTET 1095

Qy       1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSEGGPPGAEPQ 1152


RESULT 14
US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
;  GENERAL INFORMATION:
;    APPLICANT:  Arnaout, M. Amin
;    TITLE OF INVENTION:  CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
;    TITLE OF INVENTION:  RESPONSES WITH BETA2 INTEGRINS
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,062A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/216,081
;      FILING DATE:  21-MAR-1994
;      APPLICATION NUMBER:  07/637,830
;      FILING DATE:  04-JAN-1991
;      APPLICATION NUMBER:  07/539,842
;      FILING DATE:  18-JUN-1990
;      APPLICATION NUMBER:  07/212,573
;      FILING DATE:  28-JUN-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Freeman, John W.
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  00786/068003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1163 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-476-062A-44

  Query Match             59.1%;  Score 3473;  DB 2;  Length 1163;
  Best Local Similarity   61.3%;  Pred. No. 4.2e-285;
  Matches  692;  Conservative  138;  Mismatches  293;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
              ||||||    |: :: ||| ||||   | |||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
              |||| ||||||||||||:|||| ||||||||||   |  | |: ||||| |    | |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
              | : : ||::: || |||||||||   :|  |  ||  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
```

```
                 : ||||:||:   || ||:  : || | |:||| :: |:  | : : |||::| ||||||
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
                |||:| || | |:|||| ||  |:||| |||| ||::  | :||| ||||| ::|:|:|
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
                :|:||| |||||:|||||||||:| |||||| ||:|||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
                | |||      |||||:: : || |:||||:  : |   |||||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
                | :  |   | | |||||:||||||||||::|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
                ||| | || |||||||||| |||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556

Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
                ||  |  ||||||||||||:|| |||||||:|||||||| ||||||  |||:|  |||||::
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
                ||| |   |:| | |: |: ||| :|||  :   :  :||:: |  ::: |    :|| ||
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
                 ||||| ||   || | |||| :  : :||||   ||   | ||:|:|| |:|| |||||
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
                :||| || || ||||:||  ||| |||  ||||||| |:||||:| |:|||  |  |:||
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
                   | | | | ||||||| | :||  |  |||| |:  | |   ||  | |:||
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
                |    |||| ||| ||   :::||  ||||  || ||:::||| |||:|||| |||:|| |
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
                |||||||||| ||:||   ||||||: | |  | |  |:|||:||||| ||:|: | ||
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
                | |||  :|   :|:  :| |  | :::  |  ||||| ::| ||::|||| | | :||:
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
                | | :||| : |||||||| |  :     : :|| ||| |: ||:: |||| ||:|:||
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
                | :| ::| || |||||||:|||||||||| |||:|||||||:|| |
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143


RESULT 15
PCT-US96-01314-44
; Sequence 44, Application PC/TUS9601314
;  GENERAL INFORMATION:
;     APPLICANT:  M. Amin Arnaout
;     TITLE OF INVENTION:  METHODS FOR IDENTIFYING INTEGRIN
;     TITLE OF INVENTION:  ANTAGONISTS
;     NUMBER OF SEQUENCES:  78
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fish & Richardson P.C.
;       STREET:  225 Franklin Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  U.S.A.
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;       COMPUTER:  IBM PS/2 Model 50Z or 55SX
;       OPERATING SYSTEM:  MS-DOS (Version 5.0)
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US96/01314
;       FILING DATE:  30-JAN-96
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/380,167
;       FILING DATE:  30-JAN-95
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  John W. Freeman
;       REGISTRATION NUMBER:  29,066
;       REFERENCE/DOCKET NUMBER:  00786/267001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 542-5070
;       TELEFAX:  (617) 542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  44:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1163
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
PCT-US96-01314-44

  Query Match             59.1%;  Score 3473;  DB 5;  Length 1163;
  Best Local Similarity   61.3%;  Pred. No. 4.2e-285;
  Matches  692;  Conservative  138;  Mismatches  293;  Indels    6;  Gaps    4;

Qy          1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60
                ||||||     |: :: ||| ||||    | ||||||||:| |||| | |||| ||||:||||
Db         20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

Qy         61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120
                 |||| ||||||||||||:|||| ||||||||||  |  | |: ||||| |    |  |:
Db         80 GLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

Qy        121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180
                | : : ||::: || |||||||||   :|  |  ||  |: | ::  | |||||:| :|
Db        138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197

Qy        181 RIHFTFKEFQNNPNPRSLVKPITQLLGRTHTATGVRKVIRELLNITNGARKNAFKILIVI 240
                : ||||:||:   || ||:  : || | |:||| :: |:  | : : |||::| ||||||
Db        198 QTHFTFEEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy        241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
                |||:| || | |:|||| ||  |:||| |||| ||::  | :||| ||||| ::|:|:|
Db        258 TDGKKEGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

Qy        301 NFEALKTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLSTVGSYDWAG 360
                :|:||| |||||:|||||||||:| |||||| ||:|||||  | :||:|  |||: |:|
Db        318 DFDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377

Qy        361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAIILRNRVQSLVLGAPRYQHIGLVAMFR 420
                | |||      |||||:: : || |:||||:  : |   |||||||||||||| |   :|
Db        378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT 437

Qy        421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPL 480
                | :  |   | | |||||:||||||||||::|||||||||||||||||||||||||||
Db        438 QVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPHYYEQTRGGQVSVCPL 497

Qy        481 PRGQRARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
                ||| | || |||||||||| |||||||||||||||||||||| |||||||:||||||||
Db        498 PRGWR-RWWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVIGAPGEEENRGAVYLF 556
```

```
Qy        541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
Db        557 HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616

Qy        601 PVLRVKAIMEFNPREVARNVFECNDQVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660
Db        617 PVLWVGVSMQFIPAEIPRSAFECREQVVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy        661 YDLALDSGRPHSRAVFNETKNSTRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db        677 LDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPITLRLNF 736

Qy        721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db        737 TLVGKPLLAFRNLRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSFPGLKSLLVG 796

Qy        781 GPREFNVTVTVRNDGEDSYRTQVTFFFPLDLSYRKVSTLQNQRSQRSWRLACESASSTEV 840
Db        797 SNLELNAEVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQLRSLHLTCDSAPVG-- 854

Qy        841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLLKANVTSENNMPRTNKTEF 900
Db        855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTTF 914

Qy        901 QLELPVKYAVYMVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLGQRSLPISLVFLVP 959
Db        915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEEKESHVAMHRYQVNNLGQRDLPVSINFWVP 974

Qy        960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVNCSIAVCQRIQCDI 1019
Db        975 VELNQEAVWMDVEVSHPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034

Qy       1020 PFFGIQEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFVRSQTE 1079
Db       1035 PSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVYSQLPGQEAFMRAQTT 1094

Qy       1080 TKVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRQYKDMMSE 1128
Db       1095 TVLEKYKVHNPTPLIVGSSIGGLLLLALITAVLYKVGFFKRQYKEMMEE 1143
```

Search completed: November 9, 2004, 12:21:14
Job time : 28 secs